Separation of HH and HZ Final States Using Spin Correlations

Master Thesis at the Faculty of Physics
of
Ludwig-Maximilians-Universität München

submitted by

Celine Stauch

born in Wiesbaden

Munich, November 27th 2023

Separation von HH und HZ Endzuständen mithilfe von Spin Korrelationen

Masterarbeit der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von

Celine Stauch

geboren in Wiesbaden

München, den 27.11.2023

Abstract

The Standard Model of particle physics performs well when describing fundamental physics at a small scale. However, it is unable to account for some observed phenomena, which leads to the conclusion that the current model needs to be extended. The investigation of the Higgs boson and di-Higgs production processes offers opportunities to improve the understanding of nature further. Di-Higgs processes are immensely rare and there is a large amount of background processes which make the measurement of these processes very difficult. A prominent background for the di-Higgs production is the production of a Higgs and Z boson. These processes are kinematically very similar since the two bosons have comparable masses and both processes have close cross sections in proton-proton collisions. However, the Higgs boson is a scalar particle while the Z boson has a spin of 1. The spin of the Z boson transfers to the final state particle and ultimately impacts their direction.
To investigate the impact of the particle spin on the final states, a method based on the Ellis-Karliner angle is applied. This observable was originally used on a three-jet system consisting of massless partons and will be modified for massive particles. For this purpose, two approaches are tested on generator level simulation data. Further, methods of improving the jet selection are investigated. These selection methods show that a precise jet selection is necessary to be able to distinguish between Higgs and Z bosons using spin correlations. The results suggest that the use of a modified Ellis-Karliner angle provides an observable which is applicable for the separation of the investigated di-Higgs processes from the specified background process.

Contents

Chapter 1

Introduction

The Standard Model of particle physics provides a theory describing three out of the four fundamental forces and possible interactions between elementary particles. While the current model is very successful at describing nature to an incredible degree, there remain still many unanswered questions. These include the nature of dark matter (1), dark energy (2) and the apparent matter-antimatter asymmetry in the universe (3). In 2012, the latest addition to the Standard Model, the Higgs boson, was discovered by the ATLAS and CMS collaborations at CERN (4)(5). This massive particle was predicted by the Higgs-mechanism (6), which provides a mechanism for the generation of the masses of the weak gauge bosons, $W^{\pm}$ and Z^0, and the fermion masses. Since its discovery, the properties of the Higgs boson have been largely investigated. However, there are still further properties which have not been verified experimentally yet. These include the interaction between several Higgs-like particles, which is called self-interaction. Further, there are possible mechanisms producing two Higgs bosons in the final state. However, these processes are predicted to be immensely rare and have not been measured yet. Additionally, the trilinear self-coupling is one of the possible production mechanisms of a di-Higgs whose existence diminishes the cross section of these processes even more. Investigating the properties of the Higgs boson offers numerous opportunities to further improve the understanding of nature, as well as for searches of Physics Beyond the Standard Model. However, the large QCD background at the Large Hadron Collider (LHC) at CERN in Geneva imposes challenges on the measurements of those rare di-Higgs processes. Additionally, there are background processes resulting in similar final states as the di-Higgs processes. This work investigates the background of single Higgs production associated with a weak gauge boson, specifically the Z boson. Even though the main production mechanism of di-Higgs processes at the LHC is via gluon fusion, its cross section is immensely small and comparable to the cross section of the single Higgs production with an associated Z boson. Additionally, the Higgs and Z bosons are similar in mass. These similarities result in kinematic similarities of both processes. However, the Higgs boson was found to most likely be a spin-0 particle while the Z boson has a spin of 1. Considering a fully hadronic decay into bottom quarks, the spin of the Z boson will be transferred to its decay products and ultimately impact their direction. Such an angular correlation could be used to distinguish between both processes. This work investigates three-jet systems built from massive bosons and their decay products based on the method of the Ellis-Karliner angle. This observable was initially used on massless partons to verify the gluon spin (7). To use a similar method, this work presents a possible modification of this observable for massive particles, which will be evaluated for distinguishing the di-Higgs from its background process.

First, chapter 2 gives a brief overview of the theoretical background of particle physics needed in later sections. Additionally, an outlook on other developments in the field is provided. Further, concepts of experimental particle physics at particle colliders are briefly introduced with special focus on the LHC and the ATLAS experiment, which is an all-purpose detector at CERN. Chapter 3 establishes an overview of the usage and the basic structure of a Neural Network and introduces some of the most common challenges of using such networks. Afterwards, chapter 4 introduces the processes investigated in this work. Additionally, the basics of the Monte Carlo framework used to generate the investigated data sets are briefly presented. Further, the original method of the Ellis-Karliner angle, as well as a possible modification are described. Chapter 5 covers the analysis of the investigated processes using spin correlation. Additionally, methods to improve the jet selection and jet recombination are studied. Lastly, this chapter evaluates the impact of the proposed observable on distinguishing the processes by using a Neural Network. Chapter 6 finally summarizes the results of this work and provides an outlook on possible next steps extending the presented analysis.

Chapter 2

Physics Background

Particle physics aims to explain the very fundamental forces of the universe. The smallest building blocks of nature are called elementary particles and their properties and interactions are described by a theory called the Standard Model. While this theory is already very successful in describing physics at an elementary scale, there are still open questions such as dark matter, leading to numerous new theories to extend the Standard Model. An experimental apparatus able to collect data from high-energy collisions is required to measure the properties of particles and to perform measurements to verify new physics models. Such an experimental framework is given by the LHC and the ATLAS detector at the European Organization of Nuclear Research CERN in Geneva.
This chapter establishes an overview of the theoretical background of particle physics needed in later sections and provides an outlook on other developments in the field. Further, the concepts of experimental particle physics at colliders are explained, with special focus on the LHC and the ATLAS detector.

2.1 Theory of Particle Physics

Due to their energy and size, elementary particles exist in two physical regimes. Since they are very small, they follow the rules of quantum mechanics. Additionally, they follow the rules of relativistic mechanics, because they are highly energetic and move at speeds close to the speed of light. A theory combining both these properties is called Quantum field theory (QFT). It describes particles as excitations of an underlying field at a very small scale. These fields interact via three out of the four fundamental forces of the universe.

Relevant aspects of the theory of elementary particles for this work are discussed in the following sections. First, an overview of the theoretical background of elementary particles and their interactions is provided. Afterwards, the Standard Model is briefly introduced and an outlook on some of the unanswered questions is given. Additionally, the latest experimentally discovered elementary particle, the Higgs boson, and its purpose is introduced. Lastly, important concepts such as the particle spin and Lorentz transformations, which will be used in the analysis presented in this work, are briefly described.

2.1.1 Elementary Particles

The smallest particles without any known substructure are called elementary particles, which form all matter. They can be classified into two groups, fermions and bosons, depending on spin quantum numbers, which are given in units of $\hbar$. Each elementary particle has a corresponding antiparticle. While the majority of the quantum numbers are the same for the corresponding antiparticle, charge-like quantum numbers change their sign. Particle-antiparticle pairs can be created even in a vacuum. Further, particle-antiparticle pairs may annihilate and thereby create energy in the form of a photon, as shown in the Feynman diagram in figure 2.1 for electron-positron annihilation.

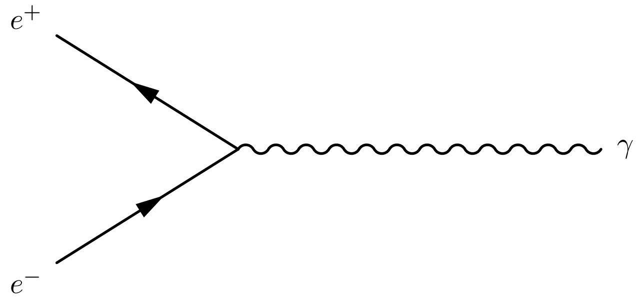


Figure 2.1: An example Feynman diagram showing a particle-antiparticle pair, represented by an electron and positron that annihilate and thereby create energy in the form of a photon.

Fermions are particles with half-integer spin while bosons have integer spin. Fermions can further be divided into leptons and quarks. Both leptons and quarks are grouped into three generations with similar masses, as shown in table 2.1. Lepton generations consist of a negatively charged lepton and a massless and chargeless neutrino.

Quark pairs are grouped by their electrical charge in up-type and down-type quarks with charges $+\frac{2}{3}$ and $-\frac{1}{3}$. Quarks cannot exist as free particles in contrast to leptons. They combine into bound states of either two or three quarks, called Hadrons. These decay into lighter particles, for instance, lighter Hadrons, which may then decay again. This process is repeated until a stable particle is produced. The phenomenon of quarks combining to build Hadrons, which is called Hadronization, was observed experimentally (8) and led to the theory of confinement (9).

	1st generation		2nd generation		3rd generation	
Quarks	u	d	c	s	t	b
Charge	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$
Mass	2.16 MeV	4.67 MeV	1.27 GeV	93.4 MeV	172.69 GeV	4.18 GeV
Leptons	e	ν_e	μ	ν_μ	τ	ν_τ
Charge	-1	0	-1	0	-1	0
Mass	0.511 MeV	< 0.8 eV	105.658 MeV	< 0.19 eV	1776.86 MeV	< 18.2 MeV

Table 2.1: The three fermion generations of leptons and quarks. The charges are given in units of the elementary charge. Quarks are represented by the respective first letter and leptons by e and the Greek letters mu μ and tau τ with their respective neutrino being indicated by an additional nu ν. All masses are given as approximations in natural units of electronvolts. Neutrinos are not completely massless which is presented by the upper bounds but are treated as such in the Standard Model. The values in this table are taken from the particle listings of the Particle Data Group (10).

The gauge bosons, which are elementary particles, are mediators of forces between particles. For simplicity, units of c = 1 and $\hbar$ = 1 are used, resulting in mass and energy in units of electronvolts (eV). Since the influence of gravity is very small this force can be neglected and only the three main forces are considered, which are electromagnetic, weak and strong

interaction.
Both the description of elementary particles and their interactions are summarized by a theory called the Standard Model of particle physics.(11)

2.1.2 The Standard Model

The Standard Model not only gives a detailed description of particles but also provides a theory on the three fundamental forces in nature: electromagnetic, weak and strong interaction. The fourth force in nature, gravity, is neglected since its influence is very small. Each interaction is described by a Lagrangian, which is constructed to reproduce the common equations of motion of the corresponding force, such as the Dirac equations. Furthermore, symmetries play an important part in physical descriptions because of their close connection to conservation laws. For example, Noether's Theorem yields conservation of angular momentum for rotational symmetry. Other symmetries, such as gauge and discrete symmetries, are relevant as well.
Each of the fundamental forces acts on different particle properties. The electromagnetic force acts on the electrical charge, the weak force acts on weak isospin and the strong force on colour charge, which is solely a property of quarks. Fields mediating these interactions are represented by gauge bosons, which act on characteristics of the respective force. The Lagrangian describes the different processes related to the force, such as a gauge boson propagating freely. Further, these Lagrangians describing the interactions need to be globally gauge invariant, which means that they are invariant under transformation performed equally at every point in space-time. Due to this global symmetry, the Lagrangian is space-time independent. Additionally, local gauge invariance is required, to allow for gauge bosons to mediate the interactions. This leads to a term in the Lagrangian which describes an interaction of a free particle with a gauge field.

Global gauge invariance can be seen by adding a divergence term $\partial_\mu M(\Phi_i, \partial_\mu \Phi^i)$ to the Lagrangian, which does not change the resulting equations of motion and therefore has no impact on the physical meaning. Such a global invariance should also hold locally. This is valid for particle physics as shown in the following. The Dirac Lagrangian

$$\mathcal{L} = i\bar{\Psi}\gamma^\mu \partial_\mu \Psi - m\bar{\Psi}\Psi \ , \tag{2.1}$$

reproduces the Dirac equation and can be modified to be invariant under a local gauge transformation of the form $\Psi \to e^{i\theta(x)}\Psi$. This is achieved by adding a term containing a gauge field A_μ which transforms like $A_\mu \to A_\mu - \partial\frac{\theta(x)}{q}$ with coupling constant q.

$$\mathcal{L} = [i\bar{\Psi}\gamma^\mu \partial_\mu \Psi - m\bar{\Psi}\Psi] - (q\bar{\Psi}\gamma^\mu \Psi)A_\mu \tag{2.2}$$

The final term of the local gauge invariant Lagrangian of equation 2.2 can be interpreted as the interaction of a fermion, which is given by its spinor Ψ, with a gauge field A_μ via a coupling constant q which affects the strength of the interaction. The first term of equation 2.2 describes the propagation of the fermion while the second term describes its mass m. Applying the same steps as shown above to an electromagnetic gauge boson field A_μ results in an expression which cannot contain mass because a mass term contains a factor $A^\mu A_\mu$, which breaks local gauge invariance. Consequently, the gauge boson has to be massless. Even though fermions can be massive, as seen in the Lagrangian in equation 2.2, they are still treated as massless. This enables splitting right and left chiral parts of the spinor. Treating these parts as separate particles may be used to differentiate between particles and antiparticles. The massless gauge boson would work for the electromagnetic and strong interaction since their respective gauge bosons, the photon and the gluon, are massless. However, the weak interaction is mediated by the $W^{\pm}$ and Z^0 bosons, which are massive and have been observed experimentally(12)(13). Combining the electromagnetic and the weak interaction to the electroweak interaction did not solve this issue. Therefore in electroweak theory, the weak gauge bosons, $W^{\pm}$ and Z^0, and the photon are created by a mixing of the massless gauge fields B and W using an angle θ. In this way, only an explanation for the generation of mass of the weak gauge bosons and the fermions is needed. A theory for that was provided by Higgs in 1964 (6), who predicted a new boson, the Higgs boson, which was used to confirm this theory in 2012.(11)(14)(15)

2.1.3 The Higgs-Mechanism

A drawback of the gauge boson Lagrangian described in chapter 2.1.2 is that it does not include the mass of the gauge bosons. Adding a mass term and applying a gauge transformation would show, that including masses in the Lagrangian breaks local gauge invariance. However experimentally it has been proven that the $W^{\pm}$ and the Z^0 are massive. To include those masses, a theory utilizing the concepts of spontaneous symmetry breaking is introduced (6). The concept of spontaneous symmetry breaking is described by a Lagrangian exhibiting a symmetry which is broken by the ground state. A simple analogy is a rotationally symmetric ball rotating on the tip of a finger: once the ball falls the rotational symmetry is broken.

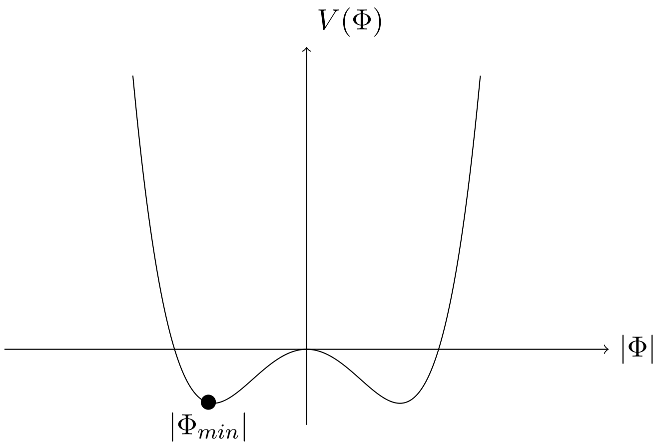


Figure 2.2: Cross section of the three dimensional Higgs potential V(Φ) of equation 2.6 showing the non-zero minima of the potential. By choosing a ground state $|\Phi|_{min}$, the breaking of the rotational symmetry of the potential is achieved.

To find a modification that fulfils the requirements above of the local gauge invariance, a potential term with a non-zero ground state is introduced to the Lagrangian. The following Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \Phi)(\partial^\mu \Phi) + \frac{1}{2}\mu^2\Phi^2 - \frac{1}{4}\lambda^2\Phi^4 \tag{2.3}$$

can be split into a "kinetic" T and "potential" V term

$$\mathcal{L} = T - V \tag{2.4}$$

in which the potential is defined as

$$V(\Phi) = -\frac{1}{2}\mu^2\Phi^2 + \frac{1}{4}\lambda^2\Phi^4 \tag{2.5}$$

This leads to a visible invariance as $\Phi \to -\Phi$. Determining the ground state for the Lagrangian of equation 2.3 requires calculating the minimum of the potential $V(\Phi)$, which is at $\Phi \neq 0$. A Lagrangian including such a potential can be interpreted as the description of a background field, which even exists in a vacuum. An example of such a field is the Higgs field, which is essential for the creation of the mass terms in the Standard Model. Writing a potential in a form similar to equation 2.5 by using a complex field $\Phi = \Phi_1 + i\Phi_2$ in three dimensions yields the following potential.

$$V(\Phi) = -\frac{1}{2}\mu^2(\Phi^\star\Phi) + \frac{1}{4}\lambda^2(\Phi^\star\Phi)^2 \tag{2.6}$$

The parameters μ and λ are not fixed and have to be determined experimentally. This is the simplest possible potential with the desired properties and it is visualized in figure 2.2. Using the above notation for the Lagrangian of equation 2.3 leads to :

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\Phi)^\star(\partial^\mu\Phi) + \frac{1}{2}\mu^2(\Phi^\star\Phi) - \frac{1}{4}\lambda^2(\Phi^\star\Phi)^2 \tag{2.7}$$

The potential in equation 2.6 has a rotational symmetry and its minima lie on a circle of radius

$$|\Phi|_{min} = \pm\frac{\mu}{\lambda}. \tag{2.8}$$

In order to use the following coordinate substitution for the real and imaginary part of Φ

$$\eta = \Phi_1 \pm \frac{\mu}{\lambda} \text{ and } \xi = \Phi_2 \tag{2.9}$$

which breaks the symmetry, one of the minima needs to be chosen as the ground state. The second field ξ can be eliminated from the new Lagrangian by applying a gauge transformation $\Phi \to \Phi'$ and allowing the resulting field Φ' to be real by demanding $\Phi_2 = 0$. This results in a Lagrangian including a mass term for the gauge bosons, as well as a new massive particle described by the field η, which is better known as the Higgs boson. While the exact shape of the Higgs potential remains unknown, the principle of employing spontaneous symmetry breaking to generate particle masses stays the same.(10)(11)

In 2012 the Higgs boson was confirmed to exist by the ATLAS and CMS experiments at CERN (4)(5). In the following years, numerous measurements were conducted that confirmed the predicted couplings and branching ratios. For instance, the coupling of the Higgs boson to the mass of fermions has already been confirmed and measured for the heavy top and bottom quarks (10). Many major decay channels have been quantified as well. It was found that the Higgs boson is most likely a spin-0 particle with even parity. Further, a measured mass at around 125 GeV implies a high probability for a meta-stable electroweak vacuum. At the LHC, which is a proton-proton collider, the most likely production mechanism is via gluon fusion with a cross section of 49 pb at the centre-of-mass energy of 13 TeV, which is rather small and makes Higgs events rare. The most dominant Higgs decays are

$H \to b\bar{b}$ with roughly 58% and $H \to WW$ with roughly 21% but there are many more decay channels (10). Currently, there are a considerable amount of measurements investigating the properties of the Higgs boson and experiments on evaluating the Higgs sector and Standard Model extensions.(10)

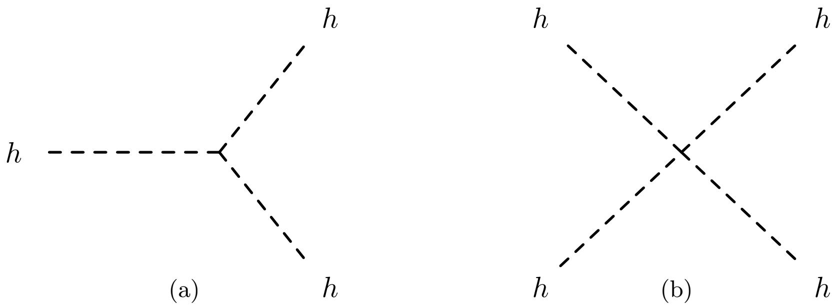


Figure 2.3: Feynman diagrams of the Higgs boson self-couplings predicted by theory. Figure2.3a is referred to as trilinear coupling and figure2.3b shows quadrolinear coupling.

Higgs self-coupling

Spontaneous symmetry breaking does not only result in a massive Higgs boson, but also in the appearance of terms in the Lagrangian that describe interactions of three or four of these newly created particles. This means, that the Higgs boson has the ability to self-couple, which implies that there are interactions between itself and other particles of its kind, which is shown in figure 2.3. The terms describing self-coupling in the Lagrangian are proportional to parameters of the potential μ and λ. Additionally, μ is proportional to the Higgs mass, which opens up the possibility of investigating the Higgs potential in more detail by measuring these parameters through self-coupling processes. Higgs self-couplings are one of the possible production channels of Higgs pairs, which have not been measured yet. Searches for Higgs pair production are important studies for probing the extensions of the Standard Model. However, the existence of Higgs self-coupling diminishes the cross section even more, which makes these processes extremely rare.(10)

2.1.4 Open Questions

While the Standard Model is verified at a high accuracy and very successful in describing physics at an elementary scale and is, there are still phenomena in nature not explainable by this theory alone. Examples of this are dark energy (2) and matter, postulated from astronomical observations (1), and the fact that the majority of the universe seems to consist of matter (3). Additionally, there are inherent deficiencies, such as the Standard Model not being able to describe the fourth fundamental force of gravity. These observations suggest that the Standard Model needs to be extended in some way to explain nature more accurately. These models are often referred to as Physics Beyond the Standard Model or New Physics. Since the current model is able to describe nature to an incredible degree, it is difficult to experimentally observe differences in this theory. This leads to a vast variety of searches for Physics Beyond the Standard Model.
Since the experimentally found values can be explained by many different models besides the Standard Model, the Higgs sector offers various opportunities to search for New Physics. Additionally, some of the decay channels have not been measured yet because of experimental challenges and it is still unclear if the Higgs boson is an elementary particle as opposed to a composite one. Measurements for the Higgs self-coupling proposed by the Standard Model

are of importance since they provide a way to measure the trilinear Higgs self-coupling, which has not been experimentally verified yet. The resonant Higgs boson is a key part of many Beyond Standard Model theories, for instance, some extensions predict its decay to potential dark matter particles, hidden valleys or dark particles. Investigating the properties of the Higgs boson further could lead to a more accurate understanding of nature.(10)

Without the guidance of the Standard Model, several new theories beyond the Standard Model have been created. One example is the Minimal Supersymmetric Standard Model which contains numerous new particles and parameters. Since these theories have not been successfully verified yet, there is no specific direction for Beyond Standard Model extensions, leading to many different theories being proposed as an extension to the current model.(16)

2.1.5 Lorentz Transformations

According to the principle of special relativity, the laws of nature take the same form in every inertial frame of reference. This also holds for the equations of the Standard Model. Time and space coordinates measured in different inertial frames of reference are related by a Lorentz transformation. This leads to the relativity of simultaneity, Lorentz contraction, time dilation and velocity addition as immediate consequences. Lorentz transformations relating frames of references moving in arbitrary directions whose axes are parallel are called boosts. A general Lorentz transformation between inertial frames of reference is a combination of a rotation and a boost. A convenient notation is given by the position-time four-vector x^μ with $\mu = 0, 1, 2, 3$:

$$x^0 = ct, x^1 = x, x^2 = y, x^3 = z \tag{2.10}$$

Writing the Lorentz transformation in terms of x^μ results in the compact form

$$x^{\mu'} = \Lambda^\mu_\nu x^\nu \tag{2.11}$$

where coefficients Λ^μ_ν are regarded as elements of a matrix Λ with $\beta = \frac{\nu}{c}$:

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.12}$$

The notation x^μ is called a contravariant four-vector. To each contravariant four-vector, there is an associated covariant four-vector x_μ, which is obtained by applying a metric $g_{\mu\nu}$ on x^ν.

$$x_\mu = g_{\mu\nu} x^\nu \tag{2.13}$$

The components of $g_{\mu\nu}$ can be stored in a matrix g

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{2.14}$$

which is its own inverse. In principle, applying a metric on the contravariant four-vector results in a covariant vector x_ν with spatial components of opposite sign.
In elementary particle physics, there is a distinction between the laboratory and the particle frame of reference. Looking at the variable of time, it is most convenient to define a proper time τ which is given by the time in the reference frame of the particle because it is invariant. Other variables, such as energy and momentum, used need to be defined in such a way, that the conservation laws hold in all inertial frames of reference. Introducing a velocity depending on the distance travelled in the lab frame and the proper time τ, the momentum is defined as

$$p^\mu = m\eta^\mu \tag{2.15}$$

where $\eta = \gamma v$ is the velocity and m the mass. The four-vector p^μ can be split into the time component

$$p^0 = \gamma mc \tag{2.16}$$

and the momentum three-vector

$$p = \gamma mv = \frac{mv}{\sqrt{1-\frac{v^2}{c^2}}} \tag{2.17}$$

with $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. Similarly, the energy is defined as

$$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} \tag{2.18}$$

which can be used to rewrite p^0 as $\frac{E}{c}$. The energy and momentum build the energy-momentum four-vector $p^\mu = (\frac{E}{c}, p_x, p_y, p_z)$. Using equation 2.15 to calculate the inner product leads to the following equation, which is invariant in all inertial frames of reference.

$$p_\mu p^\mu = \frac{E^2}{c^2} - p^2 = m^2c^2 \tag{2.19}$$

Analyzing equation 2.17 and 2.18 for mass $m = 0$, the numerator vanishes and the equations for momentum and energy become indeterminate at speeds close to the speed of light. Assuming equation 2.19 still applies, then the energy is defined as $E = |p|c$ for the massless case. This definition is useful for describing the photons and neutrinos as they are massless particles.(11)(14)

2.1.6 Spin

The particles in the Standard Model contain a quantum number called spin, which describes the intrinsic angular momentum of an elementary particle. Half-integer spin particles are fermions, which follow the Pauli exclusion principle, which states that there cannot be two identical fermions simultaneously with the same quantum numbers at the same spot. The bosons, which are integer spin particles, do not have such a restriction and can cluster in identical states. Further, the spin of composite particles can differ from their component particle spin. For example, the ground state of an atom may have a spin of 0 and behave like

a boson while the quarks and leptons of which the atom consists are fermions.
Noether's theorem results in conservation laws that are associated with continuous symmetries of the equations of motion. For example, rotational symmetry results in the conservation of angular momentum. The tensor representing the relativistic angular momentum can be separated into an extrinsic orbital part L and an intrinsic spin part S. While the conservation of the total angular momentum has to hold, this does not have to be true for the orbital and the spin parts separately. The intrinsic angular momentum, which is referred to as spin, can be described as a rotation around an axis. Since an axis involves a direction, the spin does not only have a magnitude but also a direction. This leads to either a + or − sign depending on the direction the axis is pointing in.

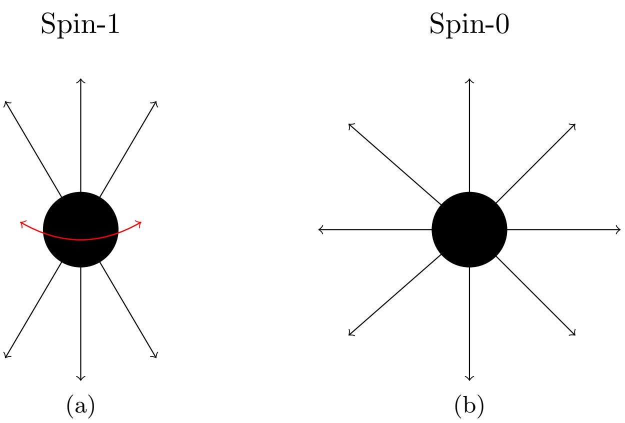


Figure 2.4: Visualization of the direction of the particle spin in the decay of a spin-1 and a spin-0 particle. The red line in figure 2.4a indicates the rotation around a vertical axis. Depending on the polarization, the spin is either pointing up or down, which is represented by the arrows of the red line symbolizing the rotation to the right or left respectively. In the case of the spin-1 particle, there is a preferred direction for the particle spin as seen in figure 2.4a. Here, the spin of the decay products points either towards or away from the direction of the spin of the initial particle as visualized by the black arrows pointing upwards or downwards. For a spin-0 particle, there is no preferred direction, as shown in figure 2.4b by the black arrows.

Considering elementary particles, the total spin needs to be conserved in a decay. In the case of a spin-1 particle, the spins of the final state needs to add up to one, which is for example realised by two spin $\frac{1}{2}$ particles. Depending on the polarization, the spin is either pointing up or down symbolizing a rotation to the right or left respectively. Therefore, there are preferred directions of spin of the decay products to achieve conservation of the total spin in case of a spin-1 particle: These are either towards or away from the direction of total spin which is visualized in figure 2.4a. Here, the red arrows indicate the rotation around a vertical axis and the black arrows point either upwards or downwards representing the spin of the decay products. For a spin-0 particle, like the Higgs boson, there are no preferred directions for the decay since there is a vast amount of possibilities to achieve a spin of zero in the final state.(11)(17)

2.2 The Experimental Framework

Particle colliders provide a high-energetic framework for high-energy processes by colliding particle beams at speeds close to the speed of light. These collisions produce numerous

particles such as hadrons which will mostly decay into lighter particles. The mostly stable particles found in the final state can be investigated. Further, particles which are of interest for certain studies can be produced directly and then investigated through the particles in the final states of their various decay channels.
Several different collider experiments are located at CERN in Geneva, Switzerland. This section will focus on the LHC and the ATLAS detector. This chapter will provide an overview of the current methods and goals, past scientific achievements and future goals at CERN in general and for the ATLAS detector in detail.

2.2.1 The LHC

The LHC is a 27 km long ring accelerator and collider installed in a tunnel below CERN previously used for the Large Electron-Positron Collider LEP. In the LHC, two proton beams travel in opposite directions and intersect at various collision points around the ring. The beams are guided and focused by strong, superconductive dipole and quadrupole magnets. Within the collision points, bunches of protons collide at a centre-of-mass energy of up to 13.6 TeV (18). The output of one collision is called an event and the number of such events per second at the LHC is given by

$$N_{event} = L\sigma_{event} \tag{2.20}$$

where σ_{event} is the cross section of the event being studied and L the luminosity of the machine. To be able to measure rare events with a small cross section, high luminosities need to be achieved at the LHC. The luminosity depends only on the properties of the beam which means that high beam energies and intensities are required to achieve high luminosities.
The LHC consists of two high luminosity experiments, ATLAS and CMS, a low luminosity experiment, LHCB, and an experiment operating with ion beams, ALICE. While LHCB specialises on B-physics, CMS and ATLAS are all-purpose detectors collecting data of proton-proton collisions.(19)

In 2012 the LHC achieved its goal of discovering the Higgs boson, which was found by both the ATLAS and CMS collaborations independently (4)(5). This verified the theory that the masses of the elementary particles are generated by the Higgs field. However, this is not the only unsolved question in particle physics the LHC wants to answer. The LHC aims further to explain Physics Beyond the Standard Model, such as the nature of dark matter and the observed asymmetry of anti-matter and matter in the universe. For this purpose, CERN conducts precision measurements to search for possible inconsistencies with the Standard Model. These could hint at the nature of a new theory needed to explain the unanswered questions and further understanding of the Higgs boson. Additionally, there are many analyses on Physics Beyond the Standard Model at the LHC, aiming to find new particles associated with new physics models, which help answer the open questions. A highly studied theory proposes a Minimal Supersymmetric Standard Model with a Higgs sector containing a second Higgs doublet. However after many years of research at the LHC, beyond Standard Model particles have not yet been observed. Since there is currently no verified model extending the Standard Model, there is still a large amount of various theories trying to answer the vast amount of open questions.(16)(19)
The second phase of the LHC aims to increase the luminosity significantly, which will be done by further improving the beam parameters. The High-luminosity LHC (HL-LHC) will be able to increase the amount of data acquired which is expected to make studies on rare decays easier. For the Higgs boson, the high luminosity will enable more precise results for precision measurements. Additionally, it may allow studying rare Higgs decays and even

di-Higgs production. Further, the mass reach for searches of new particles will be enlarged. In total, it is expected that the large data sets provided by HL-LHC improve the ability to detect new physics.(20)

2.2.2 The ATLAS Detector

In general, detectors are used to collect information about the particles in the final state of a collision in order to analyze them afterwards. For this purpose, they consist of different layers that serve different purposes. There are devices used for identifying particle trajectories, measuring the energy of the particle and measuring the momentum. An example of such a detector is the ATLAS detector at the LHC shown in figure 2.5 which is briefly described in the following.
The innermost detector layer is the tracking detector using a strong magnetic field and pixel detectors for momentum and vertex measurements, pattern recognition and electron identification. The next layer after the inner detector is the electromagnetic calorimeter used for precision measurements of photon and electron properties. Surrounding it is the hadronic calorimeter. Here, hadrons decay and hadronize, thereby creating lighter hadrons in the final state. These carry almost the same direction of travel as the initial particle which causes that the created particles distribute within a cone surrounding the initial particle. Processing data and estimating the properties of the initial particle is done by searching for and combining close-laying final state particles into one. The results of this procedure reproducing the initial particle is called a jet. In addition to measuring the hadronic decay products, the measurement of missing transverse energy also takes place in the hadronic calorimeter. Finally, the outermost layer is called the muon system, which is inside a magnetic field similar to the inner detector. This detector part is used for muon identification and measuring their tracks.

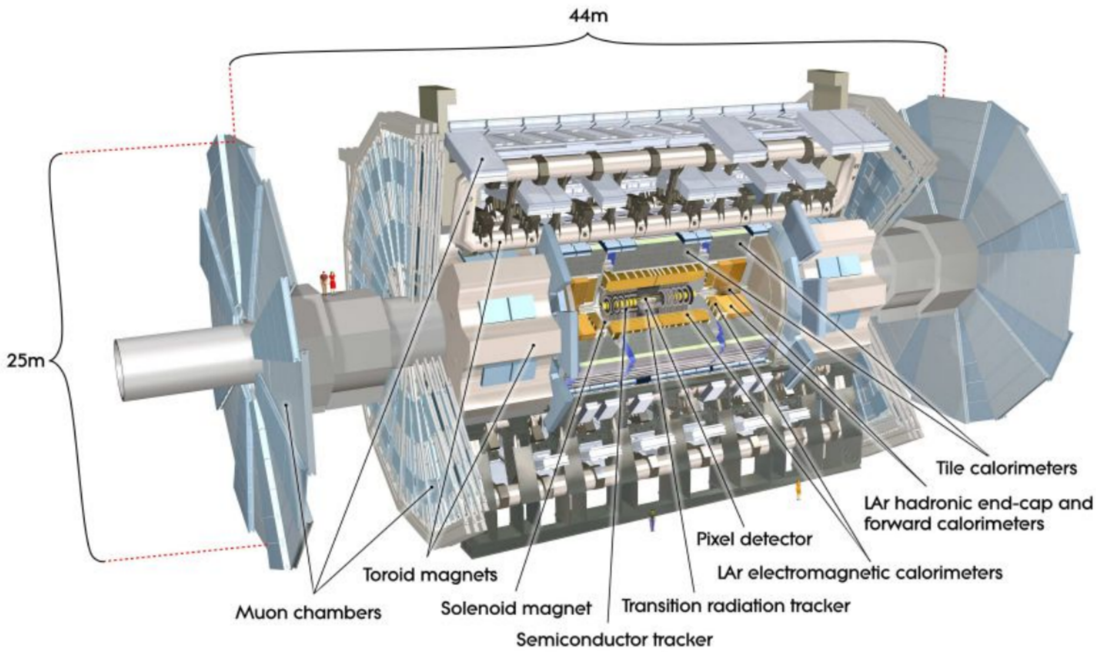


Figure 2.5: Cross section of the ATLAS detector with labels of the most important components for particle tracking and measurements.(21)

Since there is an enormous amount of data in every event far exceeding the possible data recording rate, a trigger system is applied, split into three layers. Each layer refines the decisions of the previous one and, if needed, applies additional selection criteria. The first layer, for instance, searches for high transverse-momentum muons, electrons, photons, jets and hadronically decaying τ-leptons. This ensures that collected data will most likely be of interest for current measurements and analyses.
The coordinate system used by the ATLAS detector is described as follows: The origin of the coordinate system is defined at the collision point while the z-axis points in the direction of the beam, the x-axis towards the centre of the LHC ring and the y-axis upwards which means that the x-y plane is transverse to the beam. Variables such as the transverse momentum p_T, transverse energy E_T and missing transverse energy E_T^{miss} are usually described in this plane. The azimuthal angle Φ is defined around the z-axis and the polar angle θ describes the angle from the z-axis. An additional coordinate often used is defined by the pseudorapidity η given by

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{2.21}$$

which indicates where in the detector an event occurred. Small values of the pseudorapidity η point towards events around the centre of the detector while larger values indicate a more forward direction for the event. The difference ΔR in the pseudorapidity-azimuthal plane is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\Phi^2}$.
Several of the intended measurements at CERN, such as the measurement of certain Higgs decay channels needed for its discovery, were taken into account for the design and performance goals of the ATLAS detector. However, one of the main challenges of a proton-proton collision is the large QCD background which makes the discovery of rare processes even harder. This imposes additional requirements on future improvements of this experimental framework.(21)

Chapter 3

Artificial Neural Networks

The term Artificial Neural Network defines a machine learning algorithm used for information processing. While there have already been methods and models developed at an earlier time, the use of such networks became more popular in recent years due to their capability to process large amounts of information and successfully perform given tasks. Nowadays, these information-processing structures are used in various fields, such as data analysis and particle physics. The structure of these Artificial Neural Networks is inspired by the current understanding of the human brain. This biological organ is a very complex structure built of thousands of nerve cells, called neurons, and synaptic connections, which are necessary to process immense amounts of information. The artificial counterpart aims to replicate this structure by grouping neurons into layers and connecting them. To transmit information through the network, each neuron receives and then generates an output, which will then be propagated through the layers.
This chapter establishes a brief overview of the most common components of Neural Networks used to evaluate the results of this work. After a brief motivation of the use of this technique, an introduction to the structure of such a network will be given. Afterwards, calculation methods used to achieve a good performance of the network are described. Lastly, a few of the more widely known challenges of such neural networks are introduced.

3.1 Motivation for Artificial Neural Networks

In the brain, neurons are the fundamental unit of computation and are connected in complex networks. Such an intricate structure is necessary to process the immense amounts of information and complex relations of data points. Additionally, this large network is able to establish new connections of neurons over time, improving its data-processing capabilities. This intricate biological structure inspired neural network algorithms for machine learning, aiming to reproduce their structure and information-processing abilities. Even though the neuron models for such algorithms are idealized, the principle of a network learning by changing connections between neurons stays the same.

An artificial neural network consists of neurons grouped into layers. These then generate an output transmitted to the next layer which is visualized in figure 3.1. After moving through several layers of neurons, the information is then finally given to an output layer. As indicated in figure 3.1, the information in such a network propagates from left to right without any communication between neurons in the same layer. In order to generate an output, called neuron value, the neurons need to be active. This is defined as an activation function contained by the neuron. Such a function evaluates the strength of the reaction of a neuron to

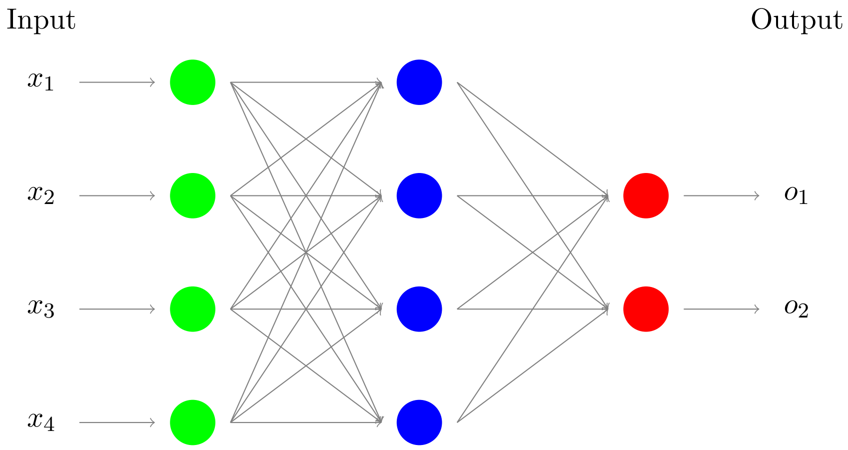


Figure 3.1: An example of a feed-forward neural network with one hidden layer. The input layer is colored green, the hidden layer blue and the output layer red. The grey arrows depict the connections between the neurons of the three layers. As indicated by the direction of the arrows, the information flow is strictly from left to right.

the incoming information. Combining these components leads to an information-processing structure which is called Neural Network (NN).(22)
While there were already various efficient learning methods and network structures developed at an earlier time, the concept of Neural Networks became prominent only in recent years through the introduction of "Deep Learning". This method, based on artificial NNs, exceeds the abilities of data-processing of a NN by employing more complex structures. Such networks are able to process immense amounts of information and in recent times are relevant for various fields, such as data analysis. Since the rise of the popularity of these techniques, the usage of properly trained NNs has demonstrated significant success due to the large abilities to capture details of given distributions.(23)

3.2 Learning in a Neural Network

NN are structures built from various parts of which each fulfils a different purpose to process large amounts of data. To be able to capture details of the input distributions, neurons are connected with each other. Each coupling differs in strength based on its importance for the performed task. To ensure a good performance of the network, an error function is introduced to compare the results of the learning process to the prediction. Further, this measurement of the NN's performance accuracy is used to improve the network by reevaluating the strength of each coupling throughout the structure. Since the goal of using a NN is to successfully perform a task, such as the classification of a target, a cross-checking procedure is applied.
This section will give a short overview of the basic components of a NN. First, introducing the general structure and building blocks of a feed-forward network. Afterwards, a closer look at the way a network improves itself is taken and lastly, a validation process is introduced.

3.2.1 Structure of a Neural Network

A NN is a structure that transmits information through a sequence of layers. Each layer consists of a number of neurons, which only feed information to the neurons of the next layer. There are no back-connections and no information is exchanged between the neurons of the same layer. Such a data processing network is called a feed-forward network. There are three different types of layers: input layers, hidden layers and output layers. The input layer on the far left of figure 3.1 receives the input data, while the output layer on the far right of figure 3.1 generates an accessible output in its neurons, which can be interpreted according to the given task. Each layer in between is called a hidden layer and its neurons are not read out and only used for computational purposes. Each connection between neurons carries a different weight, which represents the importance of the information transmitted via this connection.

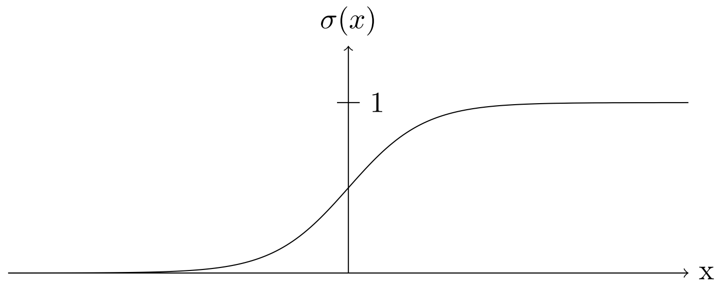


Figure 3.2: An example for a non-linear activation function defined by equation 3.1, the sigmoid function. All input values are normalized to a range of 0 to 1. Once a certain threshold is surpassed, where the values of the activation function are above zero, the neuron is active.

Every neuron in a network contains a so-called activation function. Such a function needs to be non-linear since applying a linear function in a hidden layer would not change the mapping of the input to the output. Additionally, non-linearity is necessary to solve tasks with not linearly separable input distributions. Further, these functions provide a measurement of the strength of the reaction to the input data. An example of an activation function is the sigmoid function shown in figure 3.2, which is defined as follows.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3.1}$$

This function normalizes the values on a range of 0 to 1, indicating how the neurons react to the input data. Additionally, the function is not symmetric around zero as seen in figure 3.2, which means that the output values will be positive. Another widely used activation function is called rectified liner unit, ReLU, function. Similarly to the sigmoid function, it is a non-linear function. An advantage of the ReLU function is that not all neurons are activated simultaneously which makes the ReLU function more efficient. As the network is trained via gradient descent, which is proportional to the derivative of the activation function, derivatives of simple form simplify and speed up the underlying calculations.

In the learning process of a NN, the input of each neuron is of different importance to train the network. This is represented by the weight of the connection. When receiving the input, the neuron computes a weighted average. For this purpose, all input values for that specific neuron are multiplied by their respective weights and then added up as visualized in figure 3.3. The result is then given to the activation function and the neuron value is calculated by

$$\text{output} = f(\sum_i x_i w_i) \tag{3.2}$$

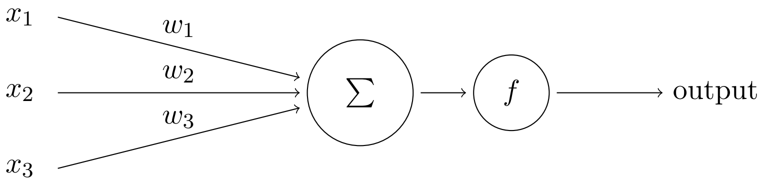


Figure 3.3: A schematic visualization of the computation chain of one neuron. The x_i with $i = 1, 2, 3$ represents the input the neuron receives with respective weights w_i, reflecting the strength of the connection. As indicated by the sum the weighted values are taken and added up. This information is afterwards given to the activation function f of the neuron and an output is generated.

Here, x_i with $i = 1, 2, 3$ represents the input values with respective weights w_i, the sum $\sum\limits_i x_i w_i$ the weighted average and f the activation function. During the learning process, all these computations are performed several times and the weights of each connection are reevaluated.(22)

3.2.2 Error Function

In a feed-forward network, the neurons are connected with each other and the information is transmitted from left to right. Each of these connections is of different strength, reflecting the importance of the transmitted information. To ensure a sufficient performance of the NN, the weights are updated over the duration of the learning process. For this purpose, error functions are introduced as a measurement of the network's accuracy. To use such an error function, a labelled data set is required. This means, that every data point has a target value t, representing the output the network should produce. These numerical labels enable a comparison of the network's results to the prediction. To generate an output as close as possible to the expected values, the error function is minimized. If the function is equal to zero its minimum is reached, meaning that the network can correctly generate outputs equal to the target value of every data point.(22)

Example: Cross-entropy loss
An example of such an error function is the cross-entropy loss defined by:

$$H = -\sum_i \left[t_i \ln f_i\right] \tag{3.3}$$

Here, t corresponds to the target values of each neuron output i and f represents the activation function used to generate the output f_i of each neuron. If $t_i = f_i$ the function is minimal and the output generated is as good as the prediction. The results of the error function can be interpreted as a probability of finding the target $t = 1$.(22)

3.2.3 Backpropagation and Gradient descent

As a measure of the network's performance accuracy, error functions are introduced that compare the output of a neuron to the prediction. In order to improve the classification of targets in the input distribution, the error of the network needs to be minimized. This minimization is then used to reevaluate the coupling strength of each neuron connection in the NN. A method minimizing the error function and recursively correcting the coupling

strengths is called gradient descent.
Each connection in the network carries a weight, which is a free parameter that represents the importance of the transmitted information. These parameters are initialized using a random distribution, which usually yields larger errors in the beginning. To improve the performance of the network, the weights are adjusted from the last to the initial layer with the aim of minimizing the error function. This method is called backpropagation.

To minimize the error function, gradient descent is applied, which minimizes the error function by adjusting the weights. For this purpose, the weights are adjusted in small steps towards the direction of the gradient. This gradient step is then repeated until the minimum is reached, which corresponds to the best possible set of parameters found during the training.
This means that the gradient of the error function H is calculated

$$\nabla H = \left(\frac{\partial H}{\partial w_1}, \frac{\partial H}{\partial w_2}, ..., \frac{\partial H}{\partial w_i} \right) \tag{3.4}$$

where H is the error function and i is the total number of weights in the network. In this method, the weight updates are defined as

$$\delta w_i = -\eta \frac{\partial H}{\partial w_i} \tag{3.5}$$

where η represents a parameter called learning rate. Since the gradient of the error function is calculated, the activation function used must be differentiable and continuous. In order to simplify and speed up the calculations during the learning process, the activation function and its gradient need to be of a simple form.(22)
The learning rate represented by η in equation 3.5 also impacts the step size of the gradient descent. While this parameter can be chosen freely, it should be relatively small to ensure that the learning process is converging.(22)

By calculating the gradients of the error function with respect to the weights, backpropagation is achieved. However, compared to the direction of the feed-forward network, the errors are updated from right to left. This backward motion through the network results from the structure of the error function and the chain rule. The equations defining this behaviour correspond to distributions that are given to the neurons. For instance, one distribution can be fed to the network and backpropagation is applied to update the weights. Iterating this step several times is called sequential training, where the number of iterations is defined as epochs.(22)

3.2.4 Validation

The backpropagation method for updating the weights of every connection of neurons through gradient descent leads to an optimization of the network by learning from its errors over time. By choosing the correct activation function and error function for the specific task, a good performance of the NN can be achieved. However, there is no proof that the network is actually learning. After many iterations, the NN adapts to the properties of the input data set which may be very specific to the given information. However, the goal of a NN is to expand from the data it has been trained on to unlabeled data sets, meaning that the more general properties of the input distributions are of interest. To ensure this generalization is

possible, the labelled data set is split into two subsets, training and validation. Since both subsets are taken from the same labelled data set, they may vary in noise but the general properties of the original data set should be present in both. Even though the network is learning by using the training subset, the error function of the validation subset is monitored as well. As long as the network is learning the general properties of the input distributions, the error functions of both subsets should decrease.
This method is useful to identify issues such as overfitting in a NN, which will be described in section 3.3. Using such a cross-validation technique to ensure proper learning of the network leads to a better generalization of the task to unknown data. Such a generalization is desirable since the goal of using a NN is to correctly classify unlabeled input data, which was, for example, obtained experimentally.(22)

3.3 Challenges of Neural Networks

A NN consists of many different components, each with its own purpose, to achieve the highest possible accuracy of predictions during the learning process. While the previously introduced structure and weight updating methods lead to a minimization of the error of the network, there still needs to be cross-checks done to validate the results. This is done, as described in section 3.2.4, by separating the labelled data set into validation and training subsets and monitoring the error function of both. This section provides an overview of the most common challenges of working with neural networks, such as underfitting, overfitting and data set bias.

Underfitting

The learning of certain properties of the input distributions is done by giving a pattern to the neurons and iterating through the layers several times, specified by the number of epochs. During the learning process, the weights are updated which means that the network is learning from its errors over time. The complexity represented by the number of neurons and hidden layers of a NN is a key factor in how well this learning process might work. If the network is very minimal and not able to sufficiently capture the properties of the input distribution it is called underfitting. This describes the challenge where the NN cannot separate two classes in the data because a lot of the details of the input distributions are missed, which is visualized in figure 3.4a. However, this may be easily fixed by changing the number of neurons and layers and thereby effectively changing the number of parameters in the network, which ultimately leads to a better performance.(24)

Overfitting

A NN can vary in complexity depending on the number of neurons in each layer and the amount of layers in the network. In the previous section, the case of a minimal network unable to sufficiently capture the details of the input distributions is discussed. However, the more widely occurring challenge is that the network learns properties of the input distributions that are specific to only the given data set. This can result in fitting noise in the training subset and is called overfitting. While underfitting can be fixed by modifying the complexity of the NN, overfitting cannot be easily fixed. Since the complexity of a NN is given by the number of parameters a network with high complexity is able to capture more details of an input distribution. This is necessary in case of intricate tasks needing a lot of information about the input data to correctly classify targets. Even though the complexity might be lowered through the usage of fewer neurons and layers, the network might not be

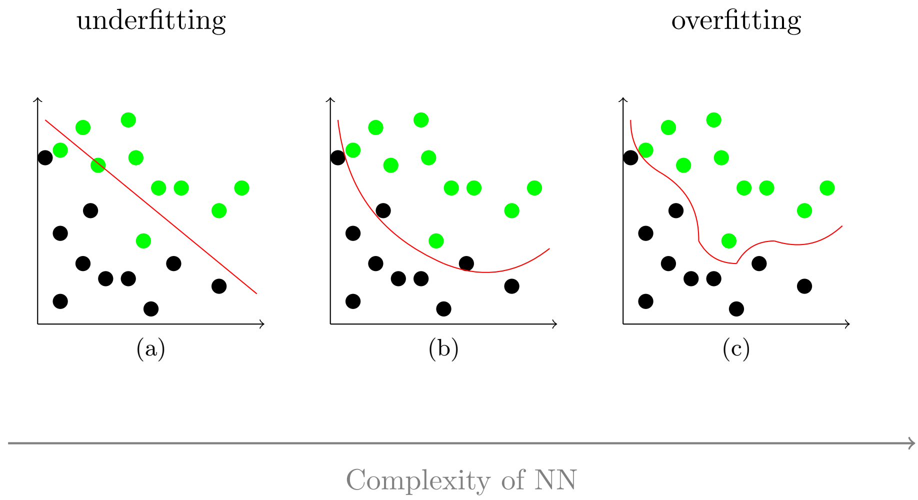


Figure 3.4: A schematic visualization of underfitting and overfitting in a network. The two classes in the data set are represented by black and green circles. The achieved separation of target classes in the input distributions is indicated by the red line. This schematic sketch visualizes how the network is fitting to the input data with increasing complexity. In figure 3.4a the case of underfitting is shown. Here, the network does not capture many details of the input distributions due to insufficient numbers of hidden layers and neurons. Therefore, it is not able to separate the two classes. Figure 3.4b indicates a good fit to the input distribution. Here, a sufficient amount of details of the input data is captured to correctly identify a large amount of data points. In figure 3.4c the case of overfitting is shown, where the NN is too complex. In this case, the network fits details specific to a given data set such as noise.

able to capture sufficient details on difficult input distributions. Due to this issue, overfitting is a widespread challenge imposed on the usage of NNs.
Since the goal of NN is to be generalized to unlabeled data, this overly specific fitting to the input data should be avoided as best as possible. For this purpose, the validation subsets, introduced in chapter 3.2.4, are monitored. Once the error function of the validation subset starts to saturate, as visualized in figure 3.5, the network is overfitting. However, since overfitting is one of the main challenges of using a NN, so-called Regularisation techniques are used. These are implemented into the network and are able to overcome the effects of overfitting. Examples of these techniques are batch normalization and dropout.(22)(24)

Pre-processing of the Input Data

The complexity of the network is not the only challenge for a NN to learn correctly. While there are methods of optimizing the weights throughout the training process of the network and achieving a high classification accuracy, it might be useful to process the input data before evaluation with the NN.

Firstly, the data can be shifted such that the mean $\langle x_k \rangle$ vanishes since large mean values may cause steep gradients in the error function for which gradient descent has difficulties functioning properly. Further, high variance and mean may lead to large values of the activation function, meaning that the neuron would react more to this input compared to smaller values. To ensure that the neuron reacts to the input with respect to its importance, the

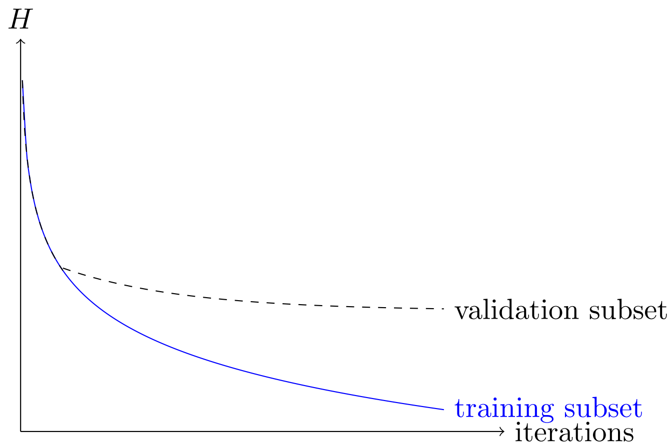


Figure 3.5: A schematic graph of the error function in case of overfitting. The data set is split into training and validation subsets, which are both monitored during the learning process of the NN. While the training subset still shows the network learning, represented by the decreasing error function, the validation subset appears to behave differently. As seen in the dashed line, the function saturates and stops decreasing. This indicates that the network fits to the noise of the input data instead of learning general properties.

input data is normalized so that the variance is equal to unity, $\sigma^2 = 1$. Additionally, a large mean of the input data leads to difficulties for the network to differentiate. In total, it is advisable to apply such a shifting and scaling method to the input data. However, this needs to be applied to every other data set the network is supposed to classify.
To ensure the NN is learning the necessary properties for the task at hand, it needs to frequently process unfamiliar data. In sequential training, the steps of receiving input data and transmitting the information through every layer, besides the output layer, are iterated several times. For the NN to receive unfamiliar data in these iterations, the data set is shuffled so that the sequence of input patterns given with every epoch is random.(22)

Data Set Bias

This chapter already introduced many methods used to ensure a NN is learning as best as possible and achieving high accuracy in classifying targets found in the data set. As discussed in the previous section, several techniques may be used to prepare the data in order to achieve better training of the network. While this is already successful in improving the performance accuracy, the impact of the given data set may need to be investigated further.

Since a NN trains based on patterns in the input distribution, changes in the input data may impact the classification accuracy. For instance, the labels of the target classes are set before handing the data set to the network to validate the obtained output. However, if these labels are incorrect, the patterns in the input distribution differ, leading to the NN training based on inaccurate information, which ultimately impacts the obtained classification accuracy. This phenomenon is called bias. While there are different types, the above example hints at specific bias in a data set. Even though the given example may be the most apparent cause of such a data set bias, any restriction to the input data before handing it to the NN should be handled with caution. For instance, certain analysis tasks may need to restrict the data set prior to training with a NN due to the complexity of the task at hand. While this does not necessarily lead to errors impacting the performance of the network, the patterns of the input distributions are much more idealized. This will impact the classification accuracy for the task. Since these restrictions are based on subjective decisions prior to the training process, every obtained result of a NN should be investigated and interpreted individually. Further, such restrictions impose a challenge on generalizing from labelled to unlabeled data sets, meaning that data set bias should be kept as small as possible. (24)

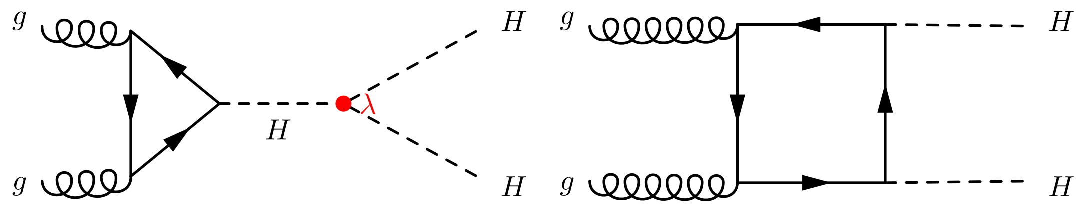
Chapter 4

Methodology

As indicated in previous chapters, the Standard Model is able to describe nature to an incredible degree. Its latest addition, the Higgs boson, provides a theory accounting for weak gauge boson and fermion masses. Further, the Lagrangian describing its interactions includes terms that describe interactions of several Higgs-like particles. Such a phenomenon is called self-coupling. However, such a coupling has not been measured experimentally yet. Additionally, Higgs bosons can be created in several ways including the production of Higgs pairs. While the single Higgs production is already very rare, the cross section for the production of a di-Higgs is even smaller. Further, the existence of trilinear self-coupling as a production process for two Higgs bosons diminishes the cross section even more. Due to its immense rarity Higgs pair production has not been measured yet. Additionally, the measurement of these processes faces more challenges due to background processes and QCD background at the LHC. An example process leading to a similar final state as the di-Higgs process is the production of a Higgs and Z boson. While there are many similarities between these processes, the spin is different for the Higgs and Z bosons. Before this work describes the analysis based on the spin properties of the investigated particles, an observable describing the impact of this property on the final state needs to be obtained. For this purpose, a method originally proposed by Ellis and Karliner to find the gluon spin, called the Ellis-Karliner angle (7), will be investigated. Further, this method will be modified to fit the desired structure. Additionally, this analysis uses Monte Carlo generated events to investigate the processes. Such simulated data opens opportunities to analyze the properties of rare events, aiming to further improve the knowledge on them and to ultimately enable experimental measurements.
This chapter gives a brief description of the two investigated processes, as a motivation of the purpose of this analysis. Additionally, an overview of the most important steps of generating the Monte Carlo data used in this work will be given. Lastly, an observable describing the impact of the particle spin on the final state in a three-jet system will be introduced and modified for investigation of the discussed processes.

4.1 Higgs-Pair Production and Background Processes

The Higgs boson is the latest addition to the Standard model and was finally discovered in 2012 (4)(5). While there have already been many measurements done on its properties since then, there are still properties suggested by its Lagrangian, which have not been validated yet. This includes the ability of the Higgs boson to couple with other particles of its type, which is called self-coupling. Since this includes the production of two Higgs bosons, which are expected to be very rare, this adds to the difficulty of measuring this property.



(a) (b)

Figure 4.1: Feynman diagrams for the dominant production mechanism of two Higgs bosons at LHC, which are produced via gluon Fusion. The process in figure 4.1a includes trilinear self-coupling of the Higgs boson with the strength depending on the coupling constant λ. In figure 4.1b the Higgs bosons are produced independently with a strength proportional to the Yukawa coupling.

The Higgs boson can be produced via different mechanisms. However, the dominating production mechanism at the LHC is via gluon Fusion (ggF). While this process has the highest cross section, the small cross section of about 49 pb (10) at 13 TeV centre-of-mass energy makes these processes rare. Measuring properties of the Higgs boson through the production of a single Higgs boson is already useful for the investigation of the Higgs potential. However, trying to measure the production of two Higgs bosons would give access to the trilinear self-coupling suggested by the Lagrangian. This property describes a Higgs boson coupling to a like-wise particle either in a trilinear or quadrolinear coupling. Similarly to the single Higgs production, there are numerous production mechanisms of which ggF has the highest cross section. Even though the ggF is the most probable production mechanism at the LHC, the cross section is about a thousand times smaller than the one for the single Higgs boson. In the Standard Model, there are two non-resonant production modes through quark loops. One includes a "box"-shaped loop of mainly top quarks, while the other one includes trilinear self-coupling. Both of these processes have a comparable amplitude. However, the diagrams shown in figure 4.1 interfere with each other negatively, leading to a smaller production rate. To measure the diagram on the right in figure 4.1 the contributions of both diagrams need to be separated.(10)

Searches for the HH production investigate measurements of the various decay channels of a Higgs boson. For a singular Higgs boson, the most prominent decay channels are the decay into a pair of bottom quarks and the decay into W bosons, as already mentioned in chapter 2.1.3. In the case of HH processes, the two Higgs bosons can decay via the same decay channel or there can be mixtures of the various existing channels. Similarly to the single Higgs production, the decay channel with the highest cross section includes Higgs bosons decaying into pairs of bottom quarks. However, searches for these processes are restricted by the immense amount of QCD background at the LHC. Additionally, the decay into bottom quarks is sensitive to the production of a vector boson and a Higgs boson. Since the Higgs boson and the weak gauge bosons are not stable, their decay channels need to be taken into account, for instance, the Z boson can decay hadronically and may even decay into a pair

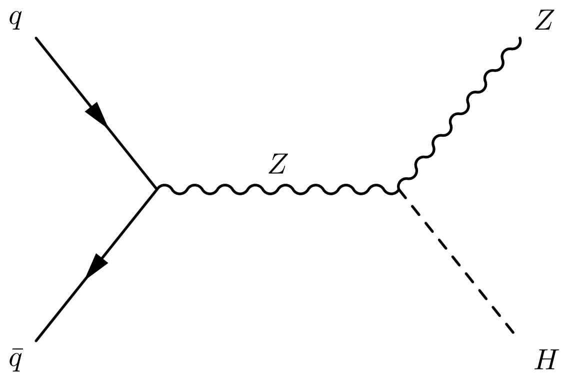


Figure 4.2: The production mechanism of a single Higgs boson with associated production of a weak gauge boson, which is the main production process for the HZ final state. The gauge boson may either be a W or Z boson. Here, the Higgs boson is radiated off of the created Z boson. This process is similar to the di-Higgs production via a self-coupling process in figure 4.1, which makes them hard to distinguish.

of bottom quarks. One of the main production mechanisms of a single Higgs boson, besides ggF, is the associated production with a gauge boson. In this production, shown in figure 4.2, the Higgs boson is radiated off of the weak gauge boson, which may either be a Z or a W boson. Even though the cross section of this process is small compared to the single Higgs production via ggF, the cross section is comparable to the one of the HH production. This imposes even more challenges on the measurements of the HH processes, since the decay into pairs of bottom quarks may happen in the HZ process, leading to a background process very similar to the desired HH process. Additional difficulties in distinguishing these processes come from the comparable masses of the Higgs and Z boson, with a Higgs mass of about $m_H = 125\,\text{GeV}$ and a Z boson mass of about $m_Z = 91\,\text{GeV}$. Since the masses of the Z boson and Higgs boson are not exact but lie on a distribution with a certain width, there are overlaps between the masses of the two bosons experimentally observed. These kinematic similarities impose many challenges on the separation of the HH production from this specific background process. However, looking at further properties of the Higgs and Z boson, which is in this work the spin, there might be additional information gained to eventually separate these processes.(10)

4.2 Generation of Monte Carlo Data Using PYTHIA

Many processes investigated in current studies are very rare, making it difficult to measure any properties of these processes and their contributing particles. Additionally, large QCD backgrounds and background processes generating similar final states impose even more challenges on these measurements. In order to increase the probabilities of finding such processes, data simulations are employed. While these are highly idealized, as they for instance ensure the process is happening in every event and suppress background processes, they still offer opportunities for analyzing the properties of the processes which then might lead to a higher chance of experimentally measuring them. For this, frameworks able to do the necessary complex computations are introduced. One of these is the POWHEG method, which combines the calculations beyond leading order with a shower generator. It has already been used for many processes including Higgs boson production associated with a vector boson. Further, it does not depend on the showering generator used. This brief introduction is focusing on the method called POWHEG BOX (25) used with PYTHIA 8 (26) for generating showering processes.

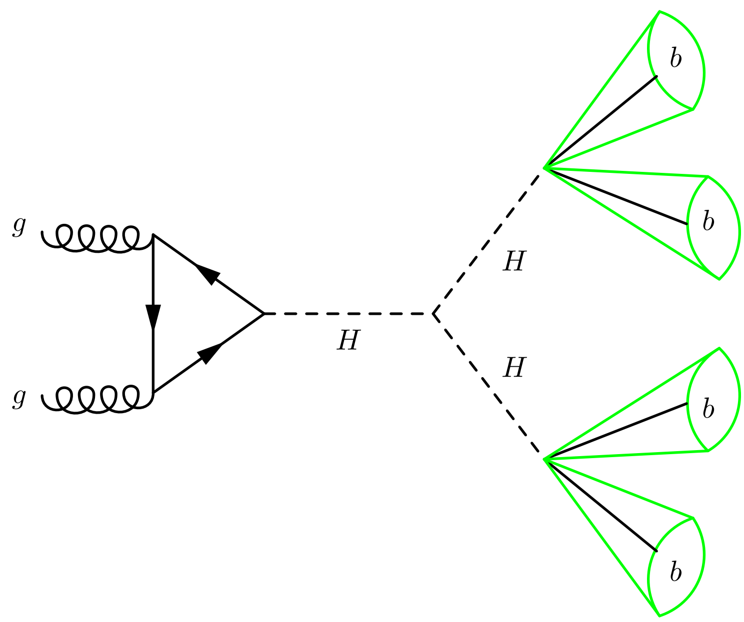


Figure 4.3: Feynman diagram of the entire process of HH production via ggF with self-coupling, which includes both Higgs bosons and their decay into the pairs of bottom quarks as a visualization for the purpose of FastJet. While the entire diagram up to the hadronization process of the quarks is described by POWHEG BOX and PYTHIA, the jets indicated as green cones are found and reconstructed by FastJet (27). Even though the chosen visualization shows cones for the resulting jets, this is not a representation of the used jet algorithm and was just drawn for simplicity.

The POWHEG BOX method, which is used to generate the HH and HZ processes discussed in this work, uses a structuring method for all incoming and outgoing particles. The particles are labelled from the incoming particles to the final state particles ordered by colour charge and mass, with indices according to Particle Data Group conventions (25). This algorithm then computes all necessary kinematic configurations to generate a process.(25) To generate the showering process, a generator such as PYTHIA needs to be interfaced with POWHEG BOX. Fortunately, POWHEG BOX does not depend on the details of the used parton shower. The most intuitive approach is to look at the characteristics of the input event and start the generation there. However, definitions of, for instance, the relative transverse momentum might differ between POWHEG BOX and PYTHIA, leading to phase-space regions being missed entirely while others are double-counted. To account for this issue, veto variables are introduced in the PYTHIA framework to ensure better matching to POWHEG BOX, which minimizes ambiguities.(28) After interfacing both code structures, PYTHIA is used to simulate particle productions in as much detail as possible. Here, a scale is used to generate the process from the hard scattering of the partons to, for instance, additional particles in the final state arising from radiation processes. As indicated above, the parton level calculations are usually imported from a separate package and only simple processes are calculated internally.
To generate a process, PYTHIA is structurally divided into three sections covering the components of an event: process level, parton level and hadron level. The process level describes the hard-scattering process, which is typically at high-energy scales. Parton level results in a representation of a realistic parton structure including jets and describing the underlying event. The last component, the hadron level, takes care of the hadronization as well as the decay of unstable hadrons, leading to an output of a realistic event as it is observed in a

detector.(26)
Lastly, the jets found in an event need to be accessible in order to analyze it. For this purpose, another tool is introduced which is used for jet-finding and analysis. The tool focused on in this work is FastJet (27). Since it is not sufficient to identify the jets visually, an algorithm projecting particles onto a jet is defined. These jet definitions can, for instance, be applied to particles. As visualized in figure 4.3 by green cones, this tool is only used after the process is generated up to the level of a real event, which includes hadronization, to reconstruct jets in the final state. Jet-finding algorithms can be classed into either sequential recombination or cone algorithms.
Sequential algorithms identify pairs of particles which are close together and then recombine them repeating that procedure until the stopping criterion is reached. The main difference between different recombination algorithms is the distance measure and stopping criterion. Cone algorithms define a specific conical region and combine the particles within this cone into a jet. These stable cones tend to be close in energy and direction to the initial parton. These algorithms differ in the strategy of finding such stable cones.(27)

The jet-finding method used in this work is a recombination algorithm, called k_T algorithm (27). It defines a distance measure between particle pairs depending on the transverse momentum of a particle with respect to the beam and a difference in the rapidity and azimuthal angle. Additionally, a jet radius defining the angular reach of the algorithm is included in the distance calculation.
There are two versions of this jet recombination algorithm: exclusive and inclusive. In the exclusive case, the smallest distance is identified. If it is the distance between particles, they are replaced with a single object and their momentum is added up. Such an object is called a PseudoJet since it is neither a particle nor a full jet. If the smallest distance found is with respect to the beam axis, then this particle is removed and contributes to the beam remnants. This is repeated until a cut value is reached after which every particle belongs to the jet. In the case of the inclusive version, there is no such cut value, the iterations continue until there are no more particles remaining. Additionally, the final jets surpassing a set transverse momentum value will be used.(27)

4.3 Spin Correlations

The HH and HZ processes introduced in chapter 4.1 are kinematically very similar. However, taking a closer look at the properties of the Higgs and Z boson shows that these particles have a different spin. As described in section 2.1.6, the particle spin may impact the direction of spin for the final state particles. To investigate this behaviour, an observable was proposed by Ellis and Karliner which describes a correlation between the spin of the initial particle and the direction of its decay products (7). This angle was originally used on three-jet systems that originate from e^+e^--collisions to confirm the spin property of the gluon. While the decay structure of the Higgs or Z boson shows a similar three-jet system, the original definition of the Ellis-Karliner angle was derived for massless partons. In order to use such an approach, this method needs to be modified to account for particle masses.
This section introduces the original method Ellis and Karliner used on a three-jet system that consists of massless partons. Afterwards, an equation based on the original Ellis-Karliner angle will be derived. This aims to include particle masses in a three-jet system to enable investigations of HH and HZ processes based on the spin properties of Higgs and Z bosons.

4.3.1 The Ellis-Karliner Angle

The spin quantum number described in chapter 2.1.6 is an intrinsic property of elementary particles. In the case of the gluon, which is described by perturbative QCD, it was predicted to be a spin-1 particle. While three hadron-jet events at the e^+e^--collider PETRA showed the desired behaviour, the spin property of the gluon still needed to be confirmed (29). For this, the directions of the three jets and their correlations were investigated, providing information about the nature of the gluon. A method using angular correlation to study the structure of these three-jet events is described by the Ellis-Karliner angle.(7)(29)

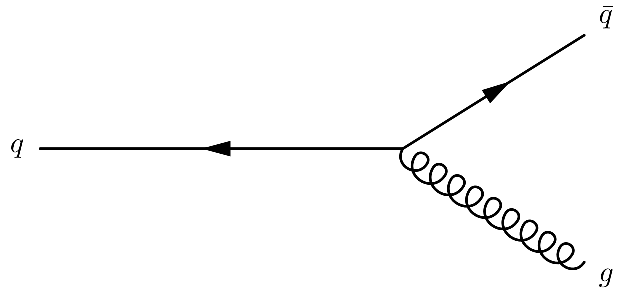


Figure 4.4: An example of the investigated three-jet structure to confirm the spin of the gluon.

The Ellis-Karliner angle is used to describe events consisting of quarks and a gluon as shown in figure 4.4. For this purpose, an axis is defined along which the highest energetic jet is aligned. After further analysis of possible three-jet events, a Lorentz boost is applied. After applying the Lorentz boost on an event of the desired structure, the two less energetic jets should be in their centre-of-mass frame and should be back-to-back. The angle θ was proposed by Ellis and Karliner to define an observable describing the angle between these jets and the previously defined axis. Calculating the distributions for the different proposed gluon spins and comparing them to data allowed to determine the gluon spin as 1.(7)(29)

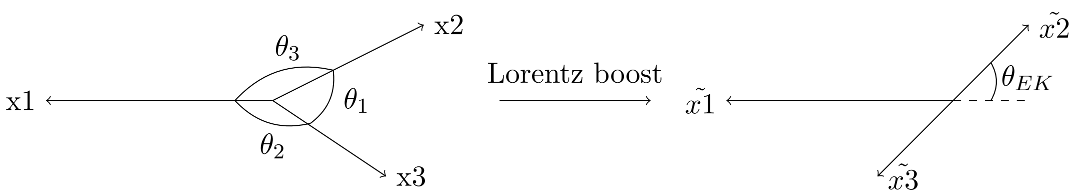


Figure 4.5: A visual representation of the three-jet structure system investigated with angular correlation. On the left, the directions of the jets are given by the angles θ_i with $i = 1, 2, 3$. According to the method proposed by Ellis and Karliner, the system is then boosted back to the rest frame of the two lower energetic jets. The variable θ_{EK}, describing the angle between these jets and the axis along the direction of the highest energetic jet, is called the Ellis-Karliner angle.

For the method proposed by Ellis and Karliner, the jets are sorted according to their energy such that

$$x_3 < x_2 < x_1 \text{ with } x_i = \frac{2E_i}{E_{cm}} \tag{4.1}$$

with E_i the energy of the quark or gluon and E_{cm} the beam energy. Afterwards, a Lorentz boost back to the centre-of-mass frame of the jets 2 and 3 is applied as shown in figure 4.5. Since the partons are massless, the observable takes the form

$$\cos\theta_{EK} = \frac{(x_2 - x_3)}{x_1} \tag{4.2}$$

where θ_{EK} is the Ellis-Karliner angle, which describes the angle between jets 1 and 2 in the rest frame of the jets 2 and 3. Defining the cross sections for a vector and scalar particle and integrating it, results in distributions of the Ellis-Karliner angle which show the differences for different proposed spin properties of the gluon.

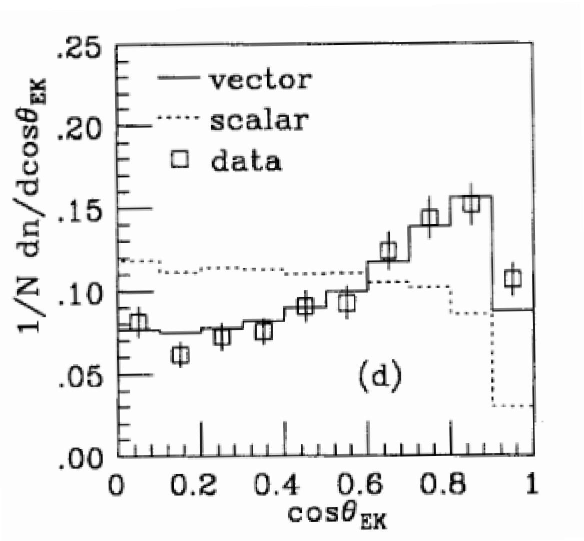


Figure 4.6: Comparison of the vector gluon described by QCD and the scalar gluon to the measured distribution. Here, the shape of the cosine of the Ellis-Karliner angle defined in equation 4.2 is visualized. As indicated in the legend, the dotted line represents the scalar gluon showing an approximately constant distribution. The solid line representing the vector gluon, which is spin-1, increases towards the right. Comparing the measurement data, visualized by the squares in the figure, to both distributions, the resulting shape looks similar to the vector gluon distribution. This comparison is taken from reference (30).

The results for investigating a measured distribution by comparison with the proposed scalar and vector gluon are shown in figure 4.6. The scalar case, represented by the dotted line, shows a rather constant distribution besides a decrease towards the far right of the interval. The shape of the proposed spin-1 gluon, which is visualized by the solid line in figure 4.6, shows an increase towards the right with a dip at the end of the interval of $\cos\theta_{EK}$. Comparing the distribution obtained from measurements to both versions, there are apparent similarities to the vector gluon case which hints towards the gluons property of spin-1.(29)(30)

4.3.2 The Modified Ellis-Karliner Angle for Jet Systems

As described in section 4.1, there are many challenges in the measurement of the HH processes. While the most apparent one is the small cross section of these processes, further difficulties arise from the background of an event. Large QCD backgrounds at the LHC as well as kinematically similar processes make the identification of HH processes much more

challenging. An example of such a background process, which will be investigated in this work, is the HZ production introduced in section 4.1. Due to the similar cross section and the close masses of the Higgs and Z bosons, these processes are hard to distinguish. However, taking a closer look at the properties of these particles, it is apparent that the spin is different. As mentioned in chapter 2.1.6 there is a correlation between the spin of a particle and the direction of the spin of the decay products. For the spin-0 Higgs boson, there is no preference in the direction of the spin while for the spin-1 Z boson a preferred direction is found. This leads to different angular correlations between the jets in the final state compared to the spin-0 particle. Such a correlation of the angle between jets and the spin of the original particle can be investigated.
An observable describing such angular correlations is defined by the Ellis-Karliner angle introduced in the previous chapter. However, this method was originally introduced for the confirmation of the gluon spin and is used on partons. Since partons are massless, equation 4.2 cannot be used on the system of a Higgs or Z boson, whose masses were mentioned in chapter 4.1, decaying into a pair of quarks. Nevertheless, this method may be modified to include massive particles.

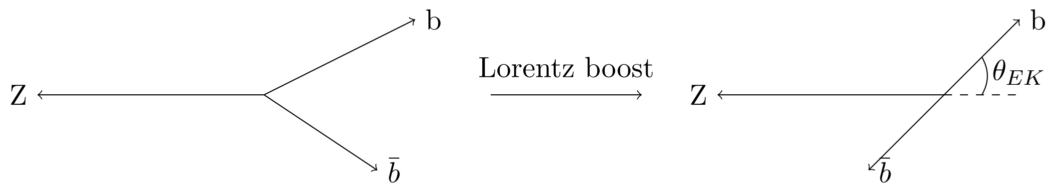


Figure 4.7: A visual representation of the three-jet system which includes the Z boson and its decay products. Similar to the previous case in figure 4.5, which includes partons, a Lorentz boost is applied so that the bottom quark pair is back-to-back. θ_{EK} indicates the desired angle based on the method of the Ellis-Karliner angle.

To use the method of the Ellis-Karliner angle, the three-jet system needs to include the Higgs or Z boson and the bottom quarks they decay to. Additionally, these systems need to be altered to account for particle masses. For a description of the modification process, this chapter focuses on the case of the Z boson presented in figure 4.7. The method to calculate the angular correlation in this new system uses a very similar approach to the original Ellis-Karliner angle. As described in the previous chapter 4.3.1, a Lorentz boost is applied to the three-jet system in such a way that the bottom quarks in figure 4.7 are back-to-back. Afterwards, the angle θ_{EK} is located between one of the lower energetic jets and the axis defined by the direction of the highest energetic jet. Looking closely at the visualization in figure 4.7, the positioning of the spin-1 particle changed compared to the original system. This is because the Z boson is the highest energetic component of the investigated system. Even though the approach resembles the original method proposed by Ellis and Karliner, the underlying calculations are more complex since equation 4.2 only accounts for the massless partons.
There are two approaches to modifying the Ellis-Karliner angle according to the system in figure 4.7. Firstly, the Lorentz boost can be calculated by using the respective functions of ROOT, which is an object-oriented framework used to analyse data (31). Secondly, a modification of the Ellis-Karliner angle may be investigated. This opens up the opportunity to derive a compact formula, which might be usable in likelihood fits. While this work will use and compare both approaches, this section will focus on the derivation of such a modified Ellis-Karliner angle.

Derivation of the modified Ellis-Karliner angle

For the derivation, the process of HZ production is split into two main parts. The first one describes the production of the Higgs and Z bosons at rest, while the second part covers the three-jet system, which includes a boson and its decay products. This derivation aims to describe the angular correlation of a three-jet system with massive particles and is based on the approach of the original Ellis-Karliner angle.

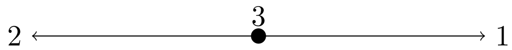


Figure 4.8: The first section of the process describes the production of the two bosons at rest. To keep a consistent naming scheme, the Z and Higgs bosons are represented by 1 and 2 respectively while the Z boson from which they originate is represented by 3. This was inspired by the original naming scheme of the investigated three-jet system (29).

Figure 4.8 visualizes the production of the Higgs and Z bosons at rest. Here, the original boson is named 3 and the Higgs and Z bosons are 2 and 1 respectively. The purpose of looking at the subsection of the process is to describe the energy and momentum of the Z boson. Since the momentum of the two bosons produced at rest is the same, it can be written as follows (10)

$$p_1 = p_2 = \frac{\sqrt{(m_3^2 - (m_1 + m_2)^2)(m_3^2 - (m_1 - m_2)^2)}}{2m_3} \tag{4.3}$$

where m_i with $i = 1, 2, 3$ represents the masses of the particles. Defining the energy of particle 1 leads to the following expression in terms of the masses. (10)

$$E_1 = \frac{m_3^2 - m_2^2 + m_1^2}{2m_3} \tag{4.4}$$

The above equations defining the energy and momentum of the system in figure 4.8 will become useful when investigating the decay into a pair of bottom quarks.

The system in figure 4.9 describes the second section of the process which covers the decay of the Z boson, 1, to the bottom quarks, q_1 and q_2. When comparing this three-jet system to

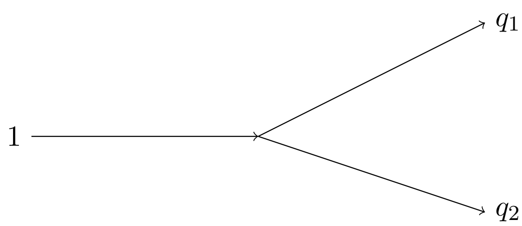


Figure 4.9: A visual representation of the three-jet system building the second subsection. Here, 1 represents the Z boson and q_1 and q_2 represent the quark pair from the decay of the Z boson. This is the system, which will be ultimately boosted to the rest frame of the quark pair.

figure 4.5, it becomes visible, that the structure is very similar, which means that if a Lorentz boost is applied, an angle comparable to the Ellis-Karliner angle should be found. To obtain an equation describing this angle, the momentum and energy of the quarks are calculated. Using conservation of energy and momentum, the following equations need to hold.

$$p_1 = p_{q_1} + p_{q_2} \tag{4.5}$$

$$E_1 = E_{q_1} + E_{q_2} \tag{4.6}$$

Additionally, the energy may be written as

$$E_{q_1} = \sqrt{m_{q_1}^2 + p_{q_1}^2} \tag{4.7}$$

where m_{q_1} is the mass of quark q_1 and p_{q_1} its momentum. For simplicity, $c = \hbar = 1$ is used. The conservation of energy may be written as follows.

$$E_1^2 = E_{q_1}^2 + E_{q_2}^2 + 2(E_{q_1} E_{q_2}) \tag{4.8}$$

Substituting E_{q_1} using equation 4.7 leads to:

$$E_1^2 = m_{q_2}^2 + p_{q_2}^2 + E_{q_1}^2 + 2(E_{q_1}\sqrt{m_{q_2}^2 + p_{q_2}^2}) \tag{4.9}$$

Further, using conservation of momentum and rewriting the momentum of q_1 in terms of mass and energy and simplifying the expression results in the following equation for the energy of quark q_1:

$$E_{q_1} = \frac{E_1^2 - m_{q_2}^2 + m_{q_1}^2}{2E_1} \tag{4.10}$$

Similarly, the energy for the second quark is obtained as:

$$E_{q_2} = \frac{E_1^2 + m_{q_2}^2 - m_{q_1}^2}{2E_1} \tag{4.11}$$

Additionally, the momentum of the quarks can be derived using a similar approach. Starting with the conservation of the momentum in equation 4.5 and rewriting one of the quark momenta using equation 4.7 leads to the following equation.

$$p_1^2 = p_{q_1}^2 + E_{q_2}^2 - m_{q_2}^2 + 2p_{q_1}\sqrt{E_{q_2}^2 - m_{q_2}^2} \tag{4.12}$$

Substituting the energy of quark q_2 with the previously obtained expression results in the following expressions for the momentum of quarks q_1 and q_2:

$$p_{q_1} = p_1 - \frac{1}{2E_1}\sqrt{(E_1^2(E_1^2 - 2m_{q_1}^2 - 2m_{q_2}^2)) + (m_{q_1}^2 - m_{q_2}^2)^2} \tag{4.13}$$

$$p_{q_2} = \frac{1}{2E_1}\sqrt{(E_1^2(E_1^2 - 2m_{q_1}^2 - 2m_{q_2}^2)) + (m_{q_1}^2 - m_{q_2}^2)^2} \tag{4.14}$$

Since energy and momentum are easily accessible from the Lorentz vectors defined for the jets in an event, these expressions will not be used for the remainder of the derivation. Additionally, it is useful to keep the final equation of $\cos\theta_{EK}$ as compact and simple as possible. However, these expressions will be used in a cross-check for the derived expression at the end of this chapter.

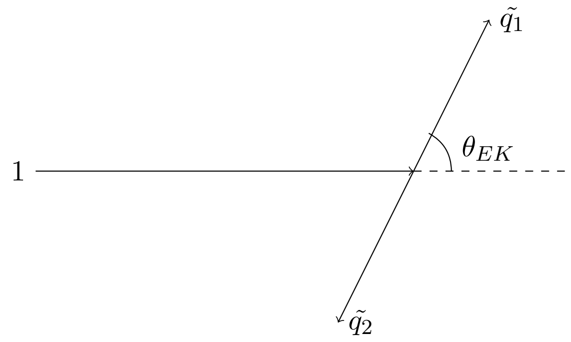


Figure 4.10: The investigated three-jet system after applying the Lorentz boost. According to the naming scheme in this work, the Z boson is represented by 1 and the quarks by $\tilde{q}_1$ and $\tilde{q}_2$ respectively. The angle θ_{EK} describes the angle between $\tilde{q}_1$ and the axis defined by the direction of 1. This is the desired observable to calculate the correlation of the Z boson's spin and the direction of its final state jets.

Boosting the three-jet system in figure 4.9 to the rest frame of the quark pair results in the quarks being back-to-back, as seen in figure 4.10. The angle indicated by $\cos\theta_{EK}$ in figure 4.10 represents the desired observable describing the correlation of the spin and the direction of the final state jets. To obtain an equation for this angle, the Lorentz transformation is calculated. Defining quantities such as the boost velocity β and the Lorentz factor γ as $\beta_1 = \frac{p_1}{E_1}$ and $\gamma_1 = \frac{E_1}{m_1}$ results in:

$$\beta_1 = -\frac{\sqrt{(m_3^2 - (m_1+m_2)^2)(m_3^2 - (m_1-m_2)^2)}}{m_3^2 - m_2^2 + m_1^2} \tag{4.15}$$

$$\gamma_1 = \frac{m_3^2 - m_2^2 + m_1^2}{2m_3 m_1} \tag{4.16}$$

The sign for β_1 depends on the direction of the Lorentz boost. Since the system is boosted back to the rest frame, the direction of the boost velocity is opposite, leading to a sign change. This may be included in the equations for the Lorentz boost. However, for simplicity in handling the expressions, the sign may be included in the boost velocity itself as well.
The Lorentz boost is generally defined by

$$p = \gamma\tilde{p} - \beta\gamma\tilde{E} \tag{4.17}$$

$$E = \gamma\tilde{E} - \beta\gamma\tilde{p} \tag{4.18}$$

which is applied to the system. Additionally, the momentum vector is projected onto the x-y plane by $\vec{p} = p\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$. The quantities γ and β of the above equations represent the Lorentz factor and the boost velocity respectively. In the case discussed in this work, the boost is applied to the rest frame, which means in the opposite direction. To distinguish between the rest frame and the laboratory frame, the values in the rest frame will be indicated by a tilde as used in figure 4.10. The boost is calculated along the longitudinal direction only, leading to the following equations for the momentum and energy:

$$\overrightarrow{p_{\tilde{q_1}}} = \begin{pmatrix} \gamma_1 p_{q_1} \cos\theta - \beta_1 \gamma_1 E_{q_1} \\ p_{q_1} \sin\theta \end{pmatrix} \tag{4.19}$$

$$E_{\tilde{q_1}} = -\beta_1 \gamma_1 p_{q_1} \cos\theta + \gamma_1 E_{q_1} \tag{4.20}$$

$$\overrightarrow{p_{\tilde{q_2}}} = \begin{pmatrix} -\gamma_1 p_{q_2} \cos\theta - \beta_1 \gamma_1 E_{q_2} \\ -p_{q_2} \sin\theta \end{pmatrix} \tag{4.21}$$

$$E_{\tilde{q_2}} = \beta_1 \gamma_1 p_{q_2} \cos\theta + \gamma_1 E_{q_2} \tag{4.22}$$

The minus signs found in the expressions for quark $\tilde{q_2}$ are due to the projection to the x-y plane. Looking at figure 4.10, the quark $\tilde{q_2}$ lies in negative direction compared to quark $\tilde{q_1}$. To project this vector onto the x-y plane, it will be multiplied by $\vec{p} = p \begin{pmatrix} -\cos\theta \\ -\sin\theta \end{pmatrix}$. Further, the expressions depend on the energy and momentum of the respective quark, as well as on the Lorentz factor and boost velocity obtained from the boson 1.

The desired angle is found in the expressions describing the Lorentz boost and represented as θ. To obtain an expression for this angle, the invariant masses are calculated. These are defined as follows:

$$M_{1/\tilde{q_1}} = m_1^2 + m_{q_1}^2 + 2(p_1 p_{\tilde{q_1}}) \tag{4.23}$$

The above expression is specifically for the case of boson 1 and quark $\tilde{q_1}$. In the last term, p_1 and $p_{\tilde{q_1}}$ describe the energy-momentum four-vectors. These can be rewritten using the definition found in section 2.1.5 as follows:

$$M^2_{1/\tilde{q_1}} = m_1^2 + m_{q_1}^2 + 2(E_1 E_{\tilde{q_1}} - |\overrightarrow{p_1}||\overrightarrow{p_{\tilde{q_1}}}| \cos\theta_{1/\tilde{q_1}}) \tag{4.24}$$

This expression depends on the energy and the absolute value of the momentum of the contributing particles, boson 1 and $\tilde{q_1}$. Additionally, a factor describing the angle between the two particle jets is represented by $\cos\theta_{1/\tilde{q_1}}$, which is the desired observable and from this point on called $\cos\theta_{EK}$. Since mass is a Lorentz invariant quantity, the particle masses are still represented by the original particles. Simplifying the above expression and dividing the invariant masses of quarks $\tilde{q_1}$ and $\tilde{q_2}$ results in:

$$\frac{M^2_{1/\tilde{q_1}}}{M^2_{1/\tilde{q_2}}} = \frac{m_1^2 + m_{q_1}^2 + 2(-\beta_1\gamma_1 p_{q_1} E_1(\cos\theta_{EK} - \cos{\theta_{EK}}^2) - \beta_1^2\gamma_1 E_1 E_{q_1} \cos\theta_{EK} + \gamma_1 E_{q_1} E_1)}{m_1^2 + m_{q_2}^2 + 2(\beta_1\gamma_1 p_{q_2} E_1(\cos\theta_{EK} - \cos{\theta_{EK}}^2) - \beta_1^2\gamma_1 E_1 E_{q_2} \cos\theta_{EK} + \gamma_1 E_{q_2} E_1)} \tag{4.25}$$

This expression only depends on quantities found in the non-boosted system in figure 4.9 as visible through the naming scheme. This was obtained by replacing the energy and momentum $E_{\tilde{q_i}}$ and $p_{\tilde{q_i}}$ for $i = 1, 2$ by using equations 4.19 to 4.22. The above expression can be rearranged in such a way, that it leads to an equation defining $\cos\theta_{EK}$. However, equation 4.25 contains terms of quadratic order in $\cos\theta_{EK}$, meaning this expression needs to be rearranged such that the quadratic formula can be applied. The final derived equation takes the following form:

$$\cos\theta_{EK} = \frac{-\left(-\frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{\beta_1 E_{q_1}}{p_{q_2}} + \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}} + \frac{\beta_1 E_{q_2}}{p_{q_2}} + 1\right)}{2\left(-1 - \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}}\right)} + \frac{\left[\left(-\frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{\beta_1 E_{q_1}}{p_{q_2}} + \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}} + \frac{\beta_1 E_{q_2}}{p_{q_2}} + 1\right)^2}{2\left(-1 - \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}}\right)}$$
$$+ \frac{4\left(1 + \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}}\right)\left(\frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{E_{q_1}}{\beta_1 p_{q_2}} + \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\left(\frac{m^2_{q_2}+m^2_1}{2\beta_1\gamma_1 E_{q_2} E_1}\right) - \left(\frac{m^2_{q_1}+m^2_1}{2\beta_1\gamma_1 E_{q_2} E_1}\right) - \frac{E_{q_2}}{\beta_1 p_{q_2}}\right)\Big]^{\frac{1}{2}}}{2\left(-1 - \frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}\frac{p_{q_1}}{p_{q_2}}\right)} \tag{4.26}$$

The above expression consists of two main parts originating from the form of the quadratic formula. The most intricate contribution to equation 4.26 stems from a square-root term, which is indicated by the term in square brackets to the power of $\frac{1}{2}$. Due to the properties of the quadratic formula, this may have either a plus or a minus sign in front. However, the positive sign found in equation 4.26 stems from the cross-check applied to the derived expression.

To validate the definition found for $\cos\theta_{EK}$, the massless limit is investigated. This should bring the three-jet system for massive particles back to the original structure of the Ellis-Karliner angle seen in figure 4.5. Further, such a limit may be investigated to validate the derived expression, as it should result in equation 4.2.

$$m_{q_1}, m_{q_2}, m_1 \longrightarrow 0 \tag{4.27}$$

The massless limit in expression 4.27 is applied by setting the masses of the three-jet structure to zero. This already leads to vanishing terms in equation 4.26. However, the energy and momentum of the quarks need to be investigated. For this purpose, equations 4.10, 4.11 and 4.13, 4.14 for the energy and momentum respectively are used. Setting the masses to zero leads to an immense simplification of these expressions. Replacing the energy and momentum of equation 4.26 and rewriting β_1 leads to an expression simplifying to

$$\cos\theta_{EK} = \frac{M^2_{1/\tilde{q}_1} - M^2_{1/\tilde{q}_2}}{M^2_{1/\tilde{q}_1} + M^2_{1/\tilde{q}_2}} \tag{4.28}$$

In the massless case, the invariant masses correspond to the particle energy. Further, the energy of boson 1 is defined as the sum of the quark energies. Comparing this to the original expression of equation 4.2 the massless limit reproduces the original definition of the Ellis-Karliner angle.

Even though the original expression of the Ellis-Karliner angle may be reproduced by applying a massless limit to the system, the derivation might still have to be verified due to its complex expressions. Additionally, it becomes apparent, that the three-jet system in figure 4.9 is in need of a more intricate description due to the particle masses. This leads to a lengthy expression for the desired observable of the correlation between particle spin and the direction of the decay products. The next chapter will investigate the derived expression as a possible differentiation between HH and HZ processes. Further, it will compare it to a different method, the application of a Lorentz boost. The goal of this derivation was to ultimately find an expression applicable in likelihood fits. The usability of the modification found for this purpose will be evaluated throughout the next chapter.

Chapter 5

Investigation of HH and HZ Processes Using Spin Correlation

To investigate the impact of the spin on distinguishing the HH and HZ final states two approaches based on the Ellis-Karliner angle are tested. For this purpose, HH and HZ processes are studied separately. This chapter describes the initial analysis of the modified Ellis-Karliner angle introduced in section 4.3.2 and an additional approach applying an explicit Lorentz transformation which is briefly mentioned in section 4.3.2. Further, methods to improve the jet selection or the HZ processes are investigated. Lastly, the achievable accuracy of distinguishing between HH and HZ final states for the selected events is tested using a NN.

5.1 Analysis of HH Processes

At the LHC, Higgs bosons are mainly produced via ggF. Even though this is the most prominent production mechanism for both single Higgs bosons and HH processes, the cross sections are rather small. This leads to Higgs boson processes being immensely rare. Additionally, the cross section of HH processes is diminished due to the destructive interference with the self-coupling process of the Higgs boson. However, investigating the properties of the Higgs boson and the trilinear self-coupling is of large interest to further the understanding of nature. While experimental measurements already face many challenges due to the rarity of HH processes, possible similarities to background processes make measurements even more difficult. An example of such a background was already introduced in chapter 4.1. This specific example imposes challenges on measuring the HH process due to the kinematic similarities. Further, the decay into bottom quarks leads to very similar final states for the investigated processes. Additionally, one of the most apparent challenges at a proton-proton collider for hadronic final states is the large QCD background. These challenges lead to a large number of analysis topics investigating properties of the Higgs boson to ultimately be able to identify such HH processes. This work focuses on the spin of the particles involved.

The Higgs boson is most likely a spin-0 particle while a Z boson has a spin of 1. This work will investigate the impact of the spin on the direction of the final state jets as a measure for distinguishing HH processes from the HZ background. For this purpose, the approach based on the Ellis-Karliner angle described in chapter 4.3.1 is used. More specifically, the three-jet system found by a boson decaying into a pair of bottom quarks is boosted back into the rest frame of the quark pair. This ultimately results in an angle between either quark jet describing the desired angular momentum correlation. While this method was originally used for massless partons to find the gluon spin, this work covers two approaches utilizing a

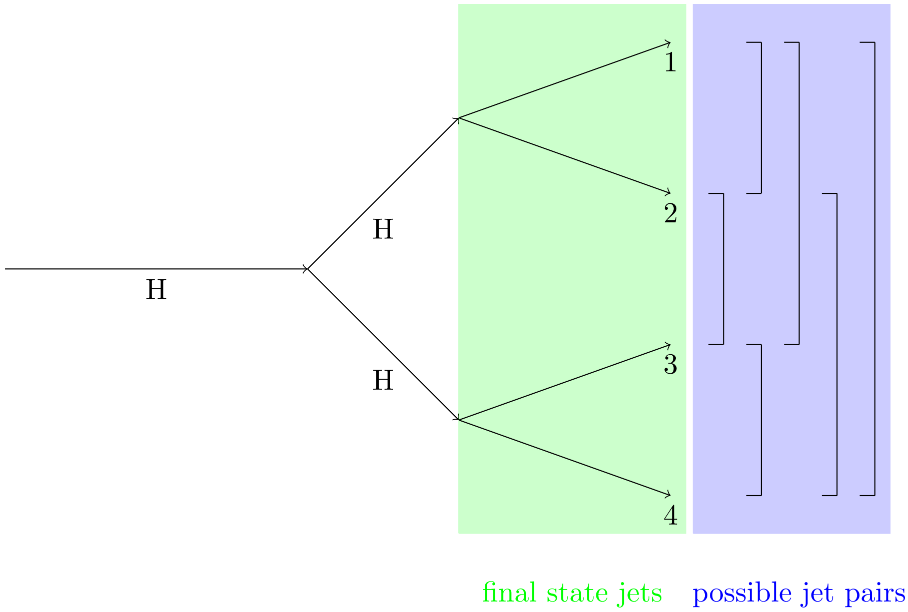


Figure 5.1: A schematic visualization of all possible combinations of the four final state jets. The jets are represented by numbers 1 to 4. Each bracket on the right indicates a possible recombination.

similar method. Firstly, in section 4.3.2, an equation describing a three-jet system of massive particles was derived defining a modified version of the Ellis-Karliner angle. Even though the derivation was focused on a system describing the Z boson decaying into quarks, the resulting expression is not specific to the Z boson but can be used for both processes. Secondly, the Lorentz boost resulting in the quarks being back-to-back may be implemented into a code structure by using built-in functions of the analysis framework ROOT. This chapter will focus on the comparison of both approaches and their challenges.

Since neither the Higgs nor the Z boson are stable particles, their decay products need to be investigated. Recombining the bosons via their decay channels yields opportunities to measure the properties of the initial particle. In the case of a Higgs boson, the most prominent decay channel is a pair of bottom quarks, which will be investigated in this work. This decay results in jets in the final state, which are identified and recombined to form the initial Higgs boson. While the decay channel with the highest branching ratio for HH processes includes decays to bottom quarks as well, this imposes an additional challenge on the measurement of these processes. Since quarks do not appear individually and undergo a process called Hadronization, the first quark initializing the jet cannot be measured. This means that the jets resulting from each boson need to be identified to recombine them. However, in an experimental measurement, it is not known which jet belongs to which decay. This work investigates Monte Carlo generated data. However, the challenge of the jet selection is still apparent. In the generated events, the jets are identified and recombined by a jet reconstruction algorithm similar to jet reconstructions in an observed event. Such algorithms only group together close-laying particles into a jet without any information on the decay they originate from. In this work, the events are generated as briefly introduced in chapter 4.2. To achieve the best possible recombination of the Higgs or Z boson, events with at least four final state jets are investigated. Further, the four jets containing the highest transverse momentum are used. These are then combined into all possible pairs of jets to find the best match for the boson masses. Figure 5.1 shows the six possible combinations of four jets, indicating each

combination by a bracket. After building these pairs, their masses are determined. Comparing those to the Higgs and Z boson mass leads to finding a combination fitting the hypothesis the best in each case. The jet reconstruction in this work uses the k_T algorithm with a radius $R = 0.4$ unless stated differently.
The selected jet pairs are then used to build the three-jet systems for the decay of the Z and Higgs bosons. Their structure is of the form described in figure 4.4 of chapter 4.3.2. The highest-energetic jet, here the boson, is aligned along the axis. The quarks build an angle to this axis as indicated in chapter 4.3.2, figure 4.9. To obtain the system that is boosted back to the rest frame for the quark pair, two approaches can be used. The first uses equation 4.26. This expression describes an implicit Lorentz boost such that the quarks are back-to-back and defines the angle between a jet and the axis. The second approach uses functions of the analysis framework ROOT. These functions use the energy-momentum four-vector of the initial boson to calculate the components of its boost velocity β. This boost is then applied to the quark jets. Lastly, the angle between the jets and the axis may be obtained by the following expression.

$$\cos\theta_{EK} = \frac{\vec{p_1}\vec{p_{q_1}}}{|\vec{p_1}||\vec{p_{q_1}}|} \tag{5.1}$$

Here, p_1 and p_{q_1} represent the momenta of the boson and one of the quarks respectively. These two approaches should lead to the same results for $\cos\theta_{EK}$. However, comparing the distributions for the Higgs and Z bosons may show differences. As described in chapter 4.3.1, a scalar particle would result in an almost constant distribution while the vector particle shows an increase towards larger values of $\cos\theta_{EK}$. However, compared to the original Ellis-Karliner angle the structure of the three-jet system changed slightly. The three-jet system investigated in this work is described in chapter 4.3.2, where the boson is the highest-energetic particle aligned with the axis. The changed placement of the spin-1 particle leads to the decreasing distribution for $\cos\theta_{EK}$.

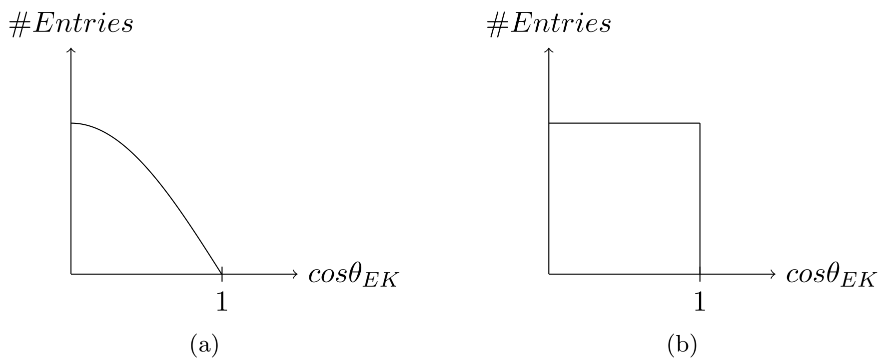
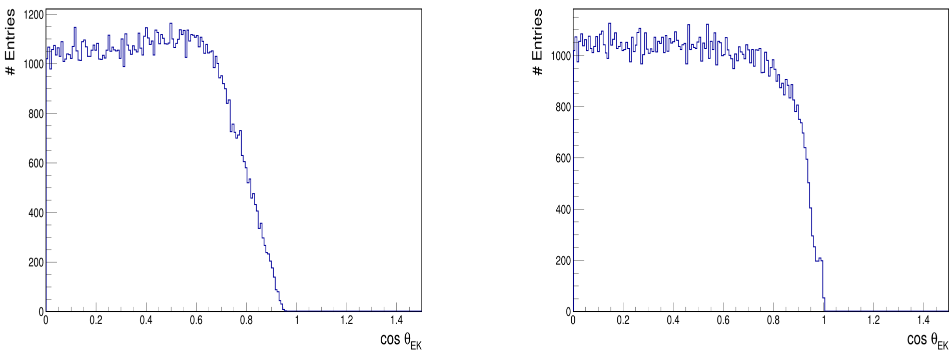


Figure 5.2: A schematic visualization of the expected distributions for $\cos\theta_{EK}$ for a scalar and a vector particle. Figure 5.2a describes the vector particle, which is in this work the Z boson. The distribution decreases towards larger values of $\cos\theta_{EK}$, showing a preference for small values of $\cos\theta_{EK}$. While the original description of the Ellis-Karliner angle showed an increase, the placement of the spin-1 particle in the used three-jet system was changed leading to the change in the distribution. Figure 5.2b represents the scalar particle, which is in this work the Higgs boson. Here, the distribution is a constant, showing that there are no preferred directions for the final state jets.

Figure 5.2 visualizes the results expected for a scalar and vector particle distribution. The sketch in figure 5.2a shows the spin-1 particle, which is in this work the Z boson, while figure 5.2b represents the scalar particle, which is the Higgs boson. These schematic distributions represent the expected results for the Higgs and Z boson in an idealized version. Further, their shapes are used to evaluate the obtained distributions during the analysis. While the previously described steps to obtain jet pairs and determine $\cos\theta_{EK}$ are used for both the HH and HZ processes, this chapter will focus solely on the HH processes.



(a) (b)

Figure 5.3: Distributions of $\cos\theta_{EK}$ for the Higgs boson obtained by MODEK and ROOTEK. Figure 5.3a shows the distribution achieved from MODEK while figure 5.3b shows the distribution using ROOTEK. Both show the desired "box" shape expected for a scalar particle. However, figure 5.3a shows an earlier decrease, leading to overall less entries.

After selecting the jet pairs that fit the boson masses best in each event, the modified Ellis-Karliner angle (MODEK) and the explicit Lorentz transformation using ROOT (ROOTEK) are calculated and a distribution for each is obtained. Figure 5.3 shows the resulting distributions without any further restrictions applied. Figure 5.3a represents the derived expression of MODEK while figure 5.3b shows the same using ROOTEK. Both distributions show a rather constant function with a "box" shape as seen for the expected distribution in figure 5.2b. However, comparing both histograms, it becomes visible that the number of entries differs. Figure 5.3b shows about 1000 Entries throughout almost the entire interval of $\cos\theta_{EK}$ with fluctuations due to statistics and a decrease only very close to $\cos\theta_{EK} = 1$, which is the end of the phase-space. However, figure 5.3a starts to decrease at around $\cos\theta_{EK} = 0.8$ and does not reach $\cos\theta_{EK} = 1$ at all. This hints at differences between both approaches. However, investigating the distributions more closely it becomes visible that figure 5.3a has fewer overall entries than figure 5.3b, hinting at issues regarding MODEK.

This behaviour is not as expected. Even though MODEK uses the Lorentz transformation implicitly to get to the derived expression, the steps taken to boost the system back to the rest frame of the quark pair are the same as in an explicit Lorentz transformation. Due to this, both approaches should yield the same results. However, the difference only occurs at higher values of $\cos\theta_{EK}$. This hints at the issues arising only towards the end of the phase-space. To investigate this further, the phase-space needs to be separated. For this, the pseudorapidity η introduced in chapter 2.1.5 is used. As described in chapter 4.1, there are two options to produce an HH using ggF. The Feynman diagrams in figure 4.1 show both production mechanisms. The process including trilinear self-coupling produces HH from a single spot. This results in the two Higgs bosons having no preferred direction they might move to. Due to this, these events are expected to be isotropic and also appearing in the centre of the detector, which means that they have small values of η. The production of HH via a box loop produces Higgs bosons at different vertices. These Higgs bosons carry information about the initial direction of the gluons. Therefore, the processes are expected to have a rather forward direction, which corresponds to larger values of η. Due to the conservation of total angular momentum, both processes may lead to similar results for the final state directions. Finding a threshold for η may aid in the evaluation, in which subset of the phase-space, the issues regarding MODEK appear.

In order to split the phase-space any value for the pseudorapidity η might be tested. However, a starting point may be given at $\eta \approx 1$ since this is near the end of the pseudorapidity region covered by the barrel calorimeters and an approximate starting point of the end-cap calorimeters (32). Additionally, an interesting value separating the phase-space is given by $\eta \approx 2.5$. This value indicates the limit of the pseudorapidity region for precision measurements.(32)(33) Testing various values of η to find a subset of the phase-space, where especially MODEK shows a distribution closer to the expectation, leads to a threshold at $\eta = 2.5$. While values above this may have slight improvements, any threshold below $\eta = 2.5$ showed a worsening of the comparability between MODEK and ROOTEK and to the expectation.

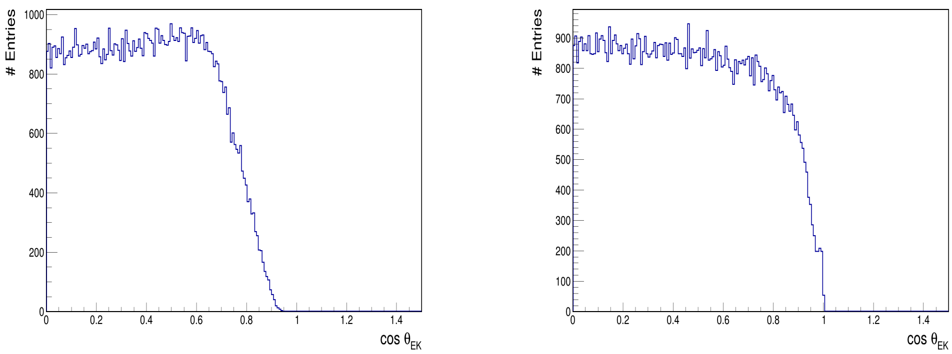
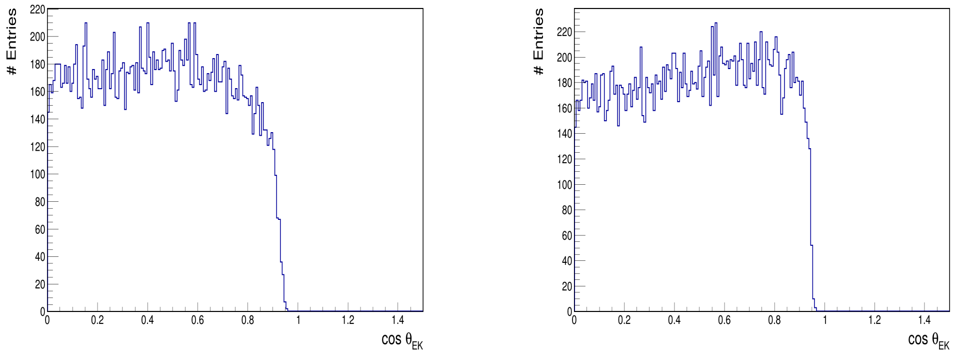


Figure 5.4: The distributions of $\cos\theta_{EK}$ in the subset of the phase-space covering the region $\eta < 2.5$. Figure 5.4a shows the result for MODEK while figure 5.4b shows the result for ROOTEK. Both show the desired "box" shape. However, figure 5.4a decreases earlier, showing no improvement in this region.

Figure 5.4 shows the results for the phase-space subset using $\eta < 2.5$, where figure 5.4a represents the MODEK distribution and figure 5.4b the ROOTEK results. Similarly to figure 5.3, both show a very "box"-like shape. However, the distribution in figure 5.4a starts decreasing much earlier than the distribution in figure 5.4b. This behaviour was already visible in the initial distributions, showing no improvement to the previous results. The main change of this pseudorapidity region is the shift in entries from 1000 in figure 5.3 to around 800 in figure 5.4.

Figure 5.5 shows the results of $\cos\theta_{EK}$ in the phase-space subset covering $\eta \geq 2.5$. Figure 5.5a shows the distribution for MODEK, here decreasing much later and slower. This leads to the shape fitting the expectations of an almost constant function better. Figure 5.5b shows the results for ROOTEK, which already fit the expected shape for the previous results. However, this distribution starts decreasing later, leading to an even better fit to the almost constant function expected. Further, both results show no entries at $\cos\theta_{EK} \approx 1$. This region of the interval is only reached at values around $\theta_{EK} = 180°$ or $\theta_{EK} = 360°$. Such cases describe a system where the quark jets are either very strongly boosted, leading to them being very close to the axis defined by the initial particle, or the jets are moving backwards according



(a)

(b)

Figure 5.5: The resulting distributions of $\cos\theta_{EK}$ in the subset of the phase-space covering $\eta \geq 2.5$. Figure 5.5a shows the result for MODEK, while figure 5.5b shows the result for ROOTEK. Further, figure 5.5a decreases considerably later compared to previous results, showing an improvement of this distribution in this region. Additionally, the results of both approaches are more equal in shape.

to the direction of the initial particle, which would require a very slow boson. However, even the strongly boosted scenario is rather rare leading to only little values effectively reaching this end-region of the phase-space. Ultimately, the distributions for both approaches show the desired "box" shape for a scalar particle. Additionally, the results of MODEK are much closer to ROOTEK and do not show as many missing entries as before. This hints at the equation derived in MODEK working more reliably in a subset of the phase-space covering large η regions. While this indicates that MODEK might work better for the HH production mechanism using the box loop, this imposes challenges for possible future applications due to precision measurements of ATLAS currently being restricted to regions of $\eta < 2.5$.

Even though the obtained results for MODEK show an improvement when separating the phase-space using $\eta \geq 2.5$, looking closely at the distributions in figure 5.5 shows that there is still a slight difference between both approaches. The results from MODEK still have fewer entries compared to ROOTEK. To find the reason for this persisting difference, the properties used in MODEK can be investigated. However, studying the behaviour of the energy and momentum of the quark jets did not lead to an explanation of this issue. Further, quantities explicitly used in MODEK and ROOTEK may be investigated and compared.

ROOTEK calculates the boost vector used to define a boost velocity β. This quantity is also found in the derived expression for MODEK, hinting at differences possibly arising from quantities used in both approaches. Figure 5.6 shows a comparison between the boost velocity obtained from ROOTEK and MODEK. MODEK defines β using information from the energy-momentum four-vector of the recombined boson, while ROOTEK uses a vector describing the boost velocity obtained by calculating the boost of the initial particle. However, since both approaches are based on a Lorentz boost to the rest frame of the quark-pair, both should result in the same boost velocity for the system. Figure 5.6 indicates this desired behaviour by showing a distribution of boost velocity pairs from MODEK and ROOTEK building a diagonal due to both approaches having the same values for β. This boost velocity is one possible example of comparable quantities between ROOTEK and MODEK. Investigating the behaviour of each variable found in both approaches always shows the desired behaviour presented in figure 5.6. The studied quantities do not show any behaviour explaining the arising differences between MODEK and ROOTEK.

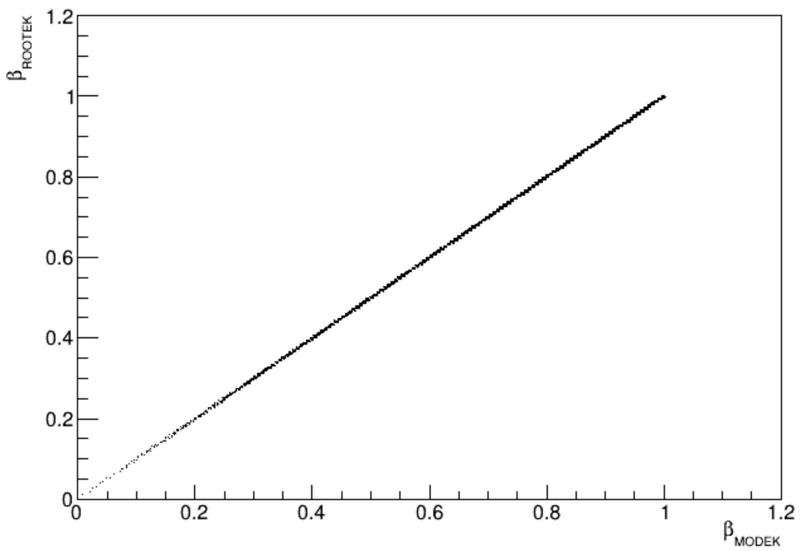


Figure 5.6: A comparison between the boost velocity used in MODEK and ROOTEK. The boost velocity obtained from MODEK is drawn on the x-axis and β from ROOTEK on the y-axis. The β pairs build a diagonal, showing that both approaches result in the same boost velocity of the system.

Further, equation 4.26 can be investigated to identify the origin of this issue. Entries not found in the given interval may be either too small or too large. However, the drawn quantity in each figure uses the absolute value of $\cos\theta_{EK}$. Additionally, a larger interval up to values of 100 has been drawn, showing no entries above $\cos\theta_{EK} = 1$. This means that the issue arises from undefined values of $\cos\theta_{EK}$. These may be caused by a division by zero or a negative value under a square-root. Investigating equation 4.26 shows that the large square-root term leads to these undefined values. While a negative value for the term in the square-root may be due to uncertainties, figure 5.7 shows that even values of around -10 may be reached. These large negative numbers are not caused by rounding errors and hint at issues regarding the expression itself.

Investigating the expression under the square-root in equation 4.26, it becomes visible that the first term is always positive while the second term may be negative. This means that there are cases where the absolute value of the second term is larger than the first term and the actual value is negative, leading to an overall negative value under the square-root. Further, studying the behaviour of the second term by investigating each contribution shows that the pre-factor defined by the invariant masses $\frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}}$ is always equal to one. This results in the following part of equation 4.26 becoming very small or even vanishing.

$$\frac{M^2_{1/\tilde{q}_1}}{M^2_{1/\tilde{q}_2}} \left(\frac{m^2_{q_2} + m^2_1}{2\beta_1\gamma_1 E_{q_2} E_1} \right) - \left(\frac{m^2_{q_1} + m^2_1}{2\beta_1\gamma_1 E_{q_2} E_1} \right) \tag{5.2}$$

This further shows that the sign of the second term under the square-root in equation 4.26 is determined by the quark energies E_{q_1} and E_{q_2}. If E_{q_1} is sufficiently smaller than E_{q_2}, the second term might result in a large enough negative value to lead to an overall negative value under the square-root. This consequently leads to undefined values for $\cos\theta_{EK}$.

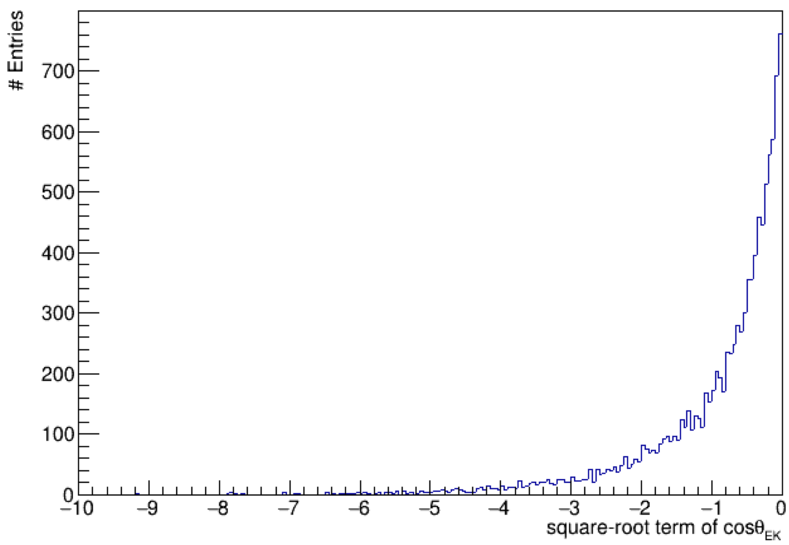


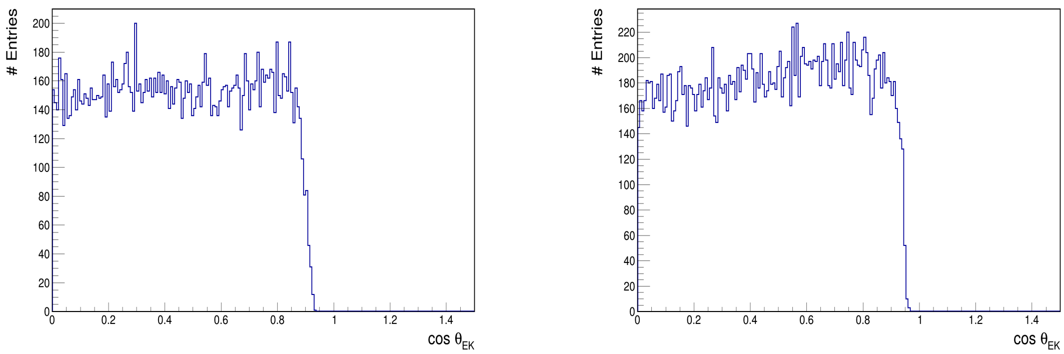
Figure 5.7: A distribution showing the range of values resulting from the term under the square-root in equation 4.26. This visualizes the cause of the undefined values for $\cos\theta_{EK}$. While values close to zero may be due to uncertainties in the calculation, larger negative values hint at issues with the derived expression itself.

While both approaches show a distribution close to the expectation for a scalar particle, the results between MODEK and ROOTEK differ. Investigating subsets of the phase-space by employing the pseudorapidity η leads to a better comparison of MODEK and ROOTEK in the region $\eta \geq 2.5$. This even leads to results closer to the expected "box" shape for both approaches. However, the results for MODEK still show fewer entries, leading to further investigation of the used quantities and the derived expression. This showed that the square-root term of the expression is negative if the energy of E_{q2} is large enough. This observation may result from challenges of the derivation, but the detailed investigation of the origin of this behaviour was beyond the scope of this work. Therefore, this work will focus on the results obtained by ROOTEK, which produced distributions fitting the expectation for a scalar particle even without using subsets of the phase-space.

5.2 Analysis of HZ Processes

A prominent background process of HH processes is the HZ production described in chapter 4.1. As already mentioned, this process may produce very similar final states to $HH \to b\bar{b}b\bar{b}$. Additionally, both processes are kinematically very similar. However, the Z boson is a spin-1 particle. This particle spin may be transferred to the final state particles, impacting their direction. Such a correlation between particle spin and the direction of their decay products is investigated using approaches based on the Ellis-Karliner angle, MODEK and ROOTEK. However, based on the results from the previous chapter, only ROOTEK will be investigated for the remainder of this work.

The HZ processes use the procedure introduced at the beginning of chapter 5.1 for the HH processes. Jet pairs are selected based on their combined mass to reconstruct Higgs and Z bosons. While the process discussed in the previous chapter only involved Higgs bosons, this chapter now needs to find both Higgs and Z bosons. For this purpose, distinct mass intervals for the Higgs and Z bosons need to be chosen. Even though these intervals may overlap due to the similar mass of the bosons, the interval chosen in this analysis excludes such an overlap by using $I_{m_Z} = [80, 100]$ GeV for the Z boson and $I_{m_H} = [115, 135]$ GeV for the Higgs boson. Additionally, a pre-sorting is applied such that the mass of the recombined Higgs boson should generally be higher than the recombined Z boson mass. While this is not necessarily the case for every Higgs and Z boson, this further helps to distinguish between the Higgs and Z bosons based on the tools used in this work. Further, the phase-space separation introduced for the HH processes will be used for a better comparison of the results of both processes.



(a) (b)

Figure 5.8: The distributions of $\cos\theta_{EK}$ for the Higgs and Z boson in the HZ processes in the phase-space subset covering $\eta \geq 2.5$. Figure 5.8a shows the distribution for the Higgs boson, which fits the expected "box" shape. Figure 5.8b shows the result for the Z boson, which looks like an almost constant distribution. However, the expected shape would show a decrease towards larger values of $\cos\theta_{EK}$.

Reconstructing the Higgs and Z boson by selecting jet pairs fitting the hypothesis best results in the distributions seen in figure 5.8. While figure 5.8a shows the expected "box"-shaped distribution for a scalar particle, figure 5.8b does not match the expectation shown in figure 5.2a. The distribution of $\cos\theta_{EK}$ for a vector particle should decrease for larger values. However, figure 5.8b shows a rather constant distribution hinting that further analysis of the system involving the Z boson is required. For this purpose, the three-jet system consisting of a Z boson and a pair of bottom quarks may be adapted to match the structure of the three-jet system originally used for the Ellis-Karliner angle. As indicated in chapter 4.3.2, the positioning of the spin-1 particle was changed due to the Z boson being the highest-energetic particle. Altering the structure such that the axis is aligned along one of the quark jets, as can be seen in figure 5.9, would result in the original form of the three-jet system used to calculate the Ellis-Karliner angle. The resulting distribution for the system on the right of figure 5.9 would then increase for large values of $\cos\theta_{EK}$, as described in chapter 4.3.1. However, the results obtained for this structure of the three-jet system shows a sharp distribution at $\cos\theta_{EK} = 1$ for the scalar and vector particle and does not lead to an improvement.

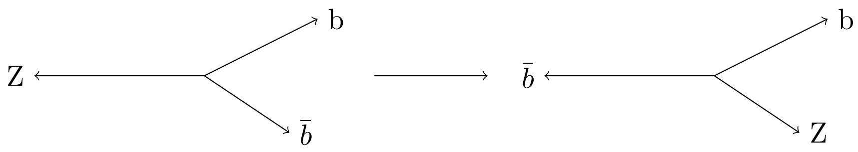
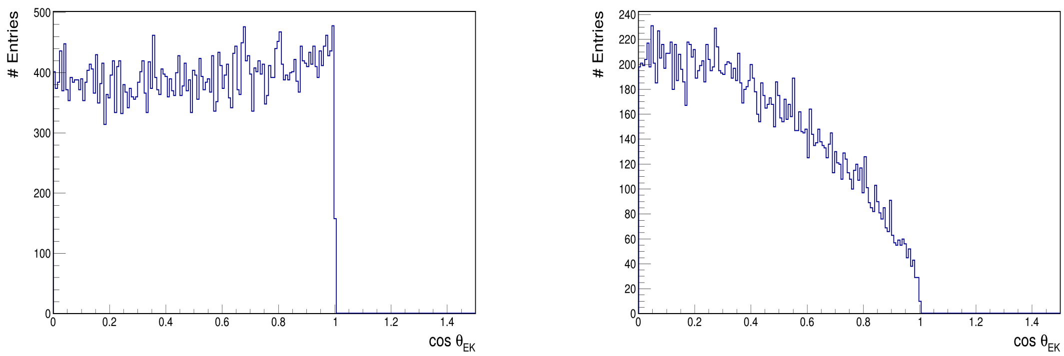


Figure 5.9: A visual representation of the change of placement of the spin-1 particle, the Z boson, in the three-jet system.

5.2.1 Investigation of the Impact of b-Tagging

Further, it needs to be investigated if the necessary information on the particle spin is available from the event generation. As described in chapter 4.2, the simulated data is obtained from a conjunction of different frameworks. In this work, POWHEG BOX interfaced with the PYTHIA showering generator was used to generate the Monte Carlo data. Definitions of quantities, such as the transverse momentum, may differ between the showering generator and POWHEG BOX. While veto variables are introduced in PYTHIA to combat this, such differences may still cause uncertainties on the transferred information. If these uncertainties are large enough, they may "wash out" the details of the quantities such as the particle spin. Further, jet reconstruction algorithms group close-laying particles together without any information on the origin of the particles, the four-vectors obtained from these jets contain large uncertainties. Each of these uncertainties might make it difficult to obtain the necessary precision to distinguish between Higgs and Z bosons. To investigate the availability of the spin information in the generation of an event, $\cos\theta_{EK}$ may be calculated for the first daughter particles of Higgs and Z bosons. For this purpose, a list of particles provided by PYTHIA is iterated to find the pairs of bottom quarks initializing the final state jets.



(a) (b)

Figure 5.10: The distributions using the first daughter particles of each boson to calculate $\cos\theta_{EK}$. Figure 5.10a shows the distribution obtained for the Higgs boson. The shape fits the expectation of being almost constant for a scalar particle. Figure 5.10b shows the resulting distribution for the Z boson. On particle level, the distribution decreases, showing a preference for small values of $\cos\theta_{EK}$, which fits the expectation for a spin-1 particle.

Using this so-called "truth" information on the bottom quarks originating from either boson results in the distributions seen in figure 5.10. Figure 5.10a shows the obtained distribution for the Higgs boson, which fits the expected shape shown in figure 5.2b for a scalar particle. Additionally, figure 5.10b shows a rapid decrease starting already at small values of $\cos\theta_{EK}$. This behaviour indicates a preferred direction of the final state particles and fits the expectation for a spin-1 particle. This means that the information about the particle spin is available for each generated event. However, it seems to become less apparent when using reconstructed jets. This may indicate, that the issue arises from the uncertainties of jet reconstruction or the approach used to select jet pairs building Higgs and Z bosons.

Since Higgs bosons couple preferably to heavy objects or are produced in association with heavy quarks, it is useful to correctly identify b-hadrons at ATLAS. For this purpose, an algorithm identifying b-jets is used. Such algorithms make use of the lifetime of b-hadrons causing a displacement of its decay particle to the collision point.(34) However, this information is not provided by the generated events. To mimic this b-tagging procedure, the position of the jets is compared to the position of the bottom quarks. While such a comparison may also be done using the energy of the jets and the bottom quarks, this work uses the spatial difference $\Delta R = \sqrt{\Delta\Phi^2 + \Delta\eta^2}$.

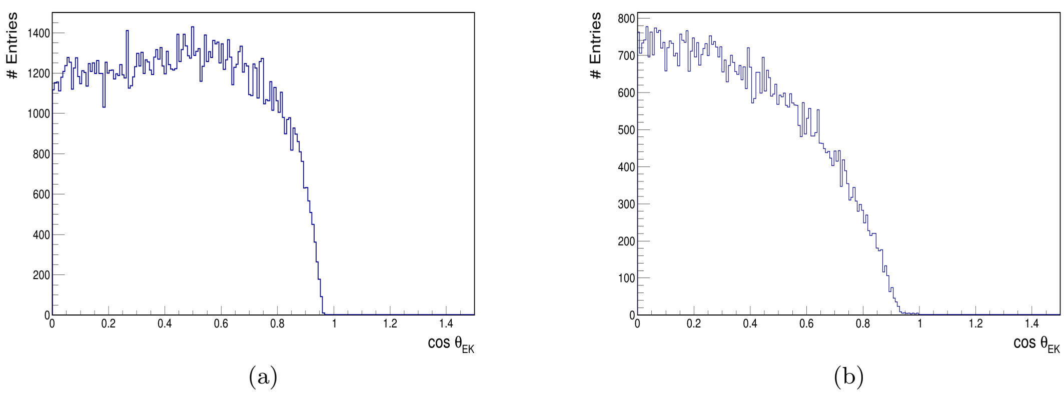


Figure 5.11: Additionally to restricting the phase-space using η, the jet selection may be improved by using the spatial difference between the bottom quark and the jet, which is given by ΔR. The distributions shown use the minimal difference to select the jet originating from a given quark. Figure 5.11a presents the results for the scalar Higgs boson and figure 5.11b the distribution for the Z boson. While the previous results already fit the expectation for a scalar particle, the result for the spin-1 particle significantly improved. This hints at better jet selections improving the achievable results for a spin-1 particle.

To find the jet fitting each bottom quark best, ΔR is calculated for every combination of jet and quark. Then, a threshold is applied and if the spatial difference of a jet is lower than this selected value, it will be set as the jet originating from the bottom quark. Depending on the value of this threshold, jets close to multiple quarks need to be excluded after being selected once. To find a measure of how well this method may work, the minimal spatial difference ΔR is investigated. Here, the difference of each of the four jets in the final state to a quark is calculated. Then, these are compared and the minimal difference is used to select the jet corresponding to the given quark. Using this approach results in the distributions shown in figure 5.11 for $\cos\theta_{EK}$. Figure 5.11a presents the results for the Higgs boson while figure 5.11b shows the distribution for the Z boson. As seen in previous results, the distribution of the scalar particle fits the expected shape in figure 5.2b, even though there is a decrease at the end of the phase-space. However, figure 5.11b shows a significant improvement to results seen in figure 5.8. The expected decrease becomes much more prominent, hinting that the minimal spatial difference would select the correct jet and quark pairs. This shows that a small distance on the $\eta - \Phi$ plane most likely corresponds to an accurate jet selection. The distributions obtained use "truth" information about the origin of the quarks, which is not available in a measurement. However, investigating the spatial difference further to find a threshold for ΔR may improve the results of $\cos\theta_{EK}$ for the Z boson.

5.2.2 Mimic b-Tagging

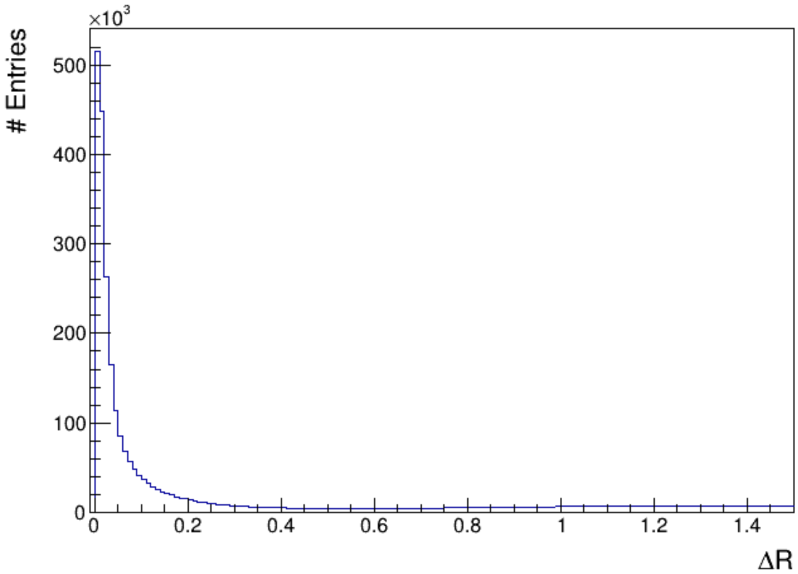


Figure 5.12: The distribution for ΔR obtained by using the minimal spatial difference between jet and quark for the jet selection. A significant amount of entries can be found at values below $\Delta R = 0.5$. However, there are many entries at larger values even reaching values above $\Delta R = 1$, which correspond to large spatial differences.

Since the distributions obtained on particle level and the minimal spatial difference show that the results expected for Higgs and Z bosons are achievable, the jet selection can be further investigated. However, using "truth" information about the bottom quarks restricts the evaluation of the ROOTEK approach for possible analyses due to restrictions of experimental measurements. In order to improve the jet selection in this work, a threshold for ΔR may be used. For this purpose, figure 5.12 shows the distribution of ΔR obtained from the distributions in figure 5.11. Here, the results for the minimal spatial difference are used to find the best possible threshold value according to a good jet selection. Investigating the distribution for ΔR in figure 5.12 shows a broad region for ΔR with a significant amount of entries at small values. Further, larger values around $\Delta R \approx 1$ describe a large distance of jet and quark in the $\eta - \Phi$ plane. The pairings selected with such large spatial differences may result in incorrect jet selections. In order to achieve sufficient jet selections, a smaller distance between the jet and the quark should be achieved. However, selecting a small threshold value may lead to low statistics for the investigated events. For further investigations, a value of $\Delta R \leq 0.5$ is chosen which ensures sufficient statistics and a good jet selection. However, as seen in figure 5.12, there are several entries for values of ΔR that are considerably larger; to improve the jet selection further, the jet substructure may be investigated.

5.2.3 Pruning of Jets

Since jets group together close-laying particles, they may be including, for instance, particles from the QCD background. These impact the properties of the resulting jets, including the jet masses. These jet masses are of importance in this work to select the jet pairs matching the Higgs and Z bosons. To remove this contamination from the jets, tools investigating the substructure of a reconstructed jet are used. Examples of such tools are: Pruning, Trimming and Mass Drop Taggers.(35) This work focuses on the use of Pruning. This tool is, for example, used to improve the reconstruction of heavy particles.(36) For this, criteria on kinematic variables are introduced to determine if a branching in the jet is likely to represent a correct reconstruction of the heavy particle. These criteria are introduced in the form of cuts that need to be passed or else, the branch is vetoed and the recombination does not occur. These cuts are represented by an angle between the daughter particles ΔR and the ratio of the minimum daughter transverse momentum and the parent transverse momentum called z.

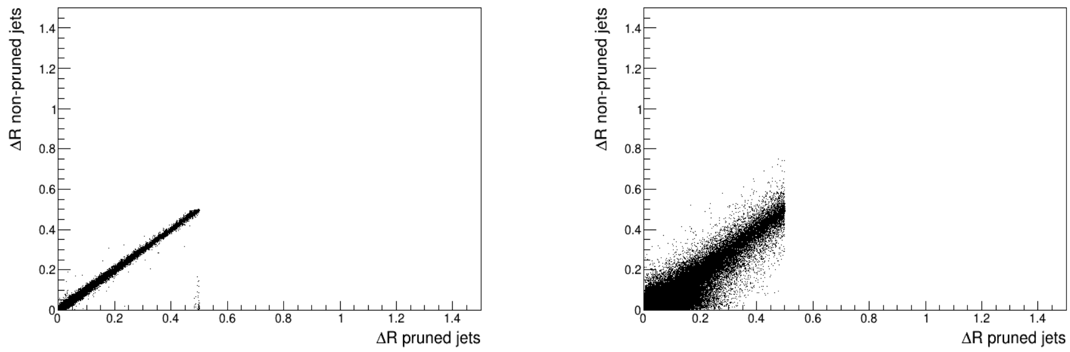
$$z \equiv \frac{min(p_{T_i}, p_{T_j})}{p_{T_{i+j}}} < z_{cut} \tag{5.3}$$

$$\Delta R_{ij} > D_{cut} \tag{5.4}$$

The expressions above define the criteria for pruning with i and j being the particles investigated at the recombination step. D_{cut} and z_{cut} define the parameters specifying the impact of the pruning and need to be selected.
Pruning is applied on reconstructed jets. For each jet, another reconstruction algorithm is applied, which may even use a different recombination algorithm than the one initially used. However, this algorithm needs to be a recombination algorithm. At every step of the recombination, the above criteria are tested. If both cuts are passed, the combination is discarded. The resulting jet is called a pruned jet, which may be compared to the initial jet.(37)(36)

The method described above may lead to a better reconstruction of the Higgs and Z boson by reducing the effects of inaccurate jet reconstructions. Further, such a tool changes the properties of the resulting jet, which may not only impact the jet masses, but help achieve a better jet selection as well. Since the four-vector describing a jet is ultimately changed after applying pruning, this may impact the spatial difference between jet and quark. To investigate the impact of pruning on the jet selection, different cut values may be tested.

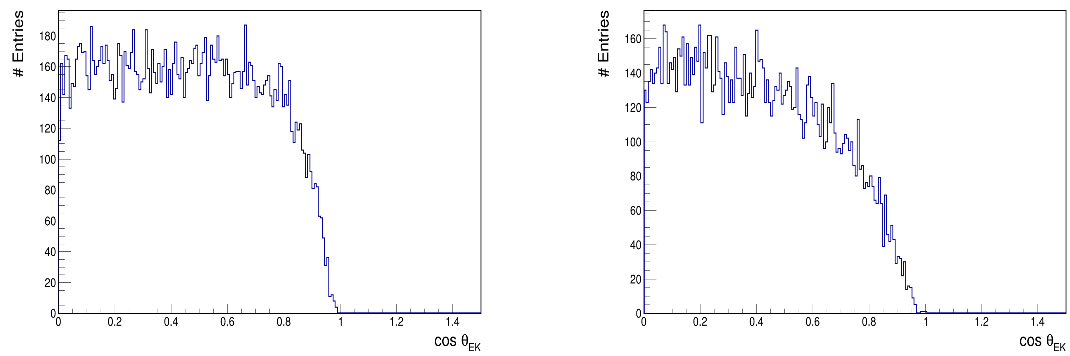
First, pruning is applied using the k_T recombination algorithm, which was used to reconstruct the initial jets. Further, the restriction $\Delta R \leq 0.5$ is used for the jet selection. In order to study the impact of different strengths of pruning, the cut values are varied. Using a small z_{cut} and large D_{cut} leads to only minimal or even no impact on the jet. However, leaving values of z_{cut} small and making D_{cut} gradually lower shows that small changes in the jet properties can be achieved. This leads to small shifts for the jet masses and small shifts in ΔR between the pruned jet and quark. While this leads to a slight improvement for the distribution of $\cos\theta_{EK}$, it is not yet producing the expected shape for the Z boson. However, using "truth" information to identify the correct jet and quark pair, it seems like this pruning is able to discard some of the branches representing an incorrect reconstruction of the Z boson. Iterating over several z_{cut} values shows that large z_{cut} results in stronger shifts in jet mass and ΔR. Combining a small D_{cut} and large z_{cut} leads to a significant shift towards lower values of ΔR. Figure 5.13 shows the impact of small pruning in figure 5.13a and a stronger pruning in figure 5.13b. As visible by comparing both figures, using $z_{cut} = 0.7$



(a) (b)

Figure 5.13: A comparison of different pruning criteria showing the impact on ΔR between jet and quark. On the x-axis, ΔR of the pruned jet is drawn and on the y-axis, ΔR of the corresponding non-pruned jet is drawn. Figure 5.13a was obtained by applying $z_{cut} = 0.05$ and $D_{cut} = 0.05$. Here, it is visible that there is no significant clustering towards lower values of ΔR. Figure 5.13b was obtained using $z_{cut} = 0.7$ and $D_{cut} = 0.05$. This resulted in a large shift in ΔR leading to a significant clustering towards smaller values.

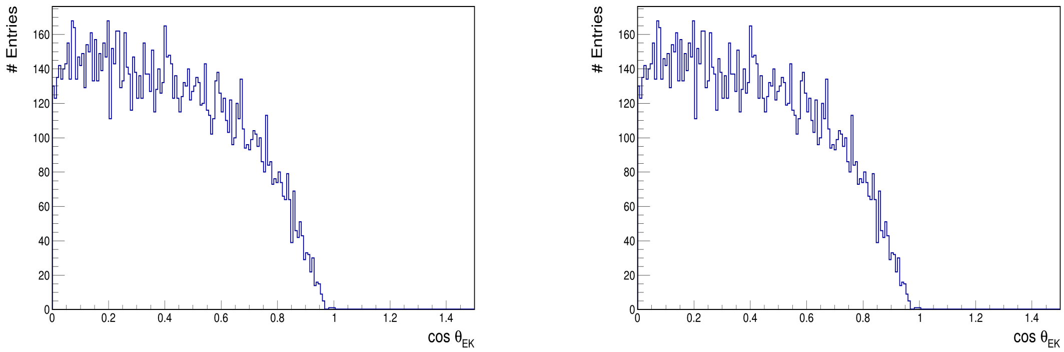
and $D_{cut} = 0.05$ significantly impacts the properties of the pruned jet. These harsh cut values result in clustering at small ΔR, which possibly leads to an improved jet selection due to the small spatial difference. Additionally, the jet masses are largely shifted to smaller values.



(a) (b)

Figure 5.14: Distributions of $\cos\theta_{EK}$ for the Higgs and Z boson using pruning with $z_{cut} = 0.7$ and $D_{cut} = 0.05$. These distributions were obtained by using $\Delta R \leq 0.5$ for the jet selection and additionally combining the jet pairs by using "truth" information on the origin of the bottom quarks. Figure 5.14a shows the results for the Higgs boson, fitting the expected shape seen in figure 5.2b. The distribution for the Z boson in figure 5.14b shows the expected decrease for larger values of $\cos\theta_{EK}$.

Since the jet masses are shifted using the strong pruning mentioned above, information on the origin of the quarks is used to obtain the distributions for $\cos\theta_{EK}$. Since "truth" information on the origin of the bottom quarks is available, the jet pairs can be selected according to the origin of the corresponding quarks. Using this information to build the three-jet system instead of the mass of the combined jets, results in the distributions shown in figure 5.14. The resulting distribution shown in figure 5.14a fits the expected "box" shape of $\cos\theta_{EK}$ for a scalar particle. Figure 5.14b shows the distribution for the Z boson, which decreases for larger values of $\cos\theta_{EK}$. This result was previously obtained on particle level and for the minimal spatial difference, showing that a more accurate jet selection is able to improve the results for $\cos\theta_{EK}$.



(a) (b)

Figure 5.15: The distributions for $\cos\theta_{EK}$ using the same approach as for figure 5.14 using the Cambridge-Aachen algorithm and anti-k_T algorithm for the recombination of the jets. Figure 5.15a shows the results using the Cambridge-Aachen algorithm and figure 5.15b was obtained using the anti-k_T algorithm. Both distributions are approximately the same and very close to the results obtained by using the k_T algorithm.

As mentioned in the introduction of pruning in section 5.2.3, the recombination algorithm used may be different to the initial jet reconstruction algorithm. For example, the Cambridge-Aachen algorithm (CA) and the anti-k_T algorithm may be used. CA is a k_T-like jet reconstruction algorithm which investigates pairwise four-vector inputs. However, this algorithm uses the transverse momentum with respect to the beam axis instead of the transverse momentum p_T (38). Further, anti-k_T uses a similar approach compared to the k_T. The difference arises from a sign change between both algorithms which leads to the anti-k_T algorithm being the inverse of the k_T algorithm (39). Since there are small differences in the distance measures used in the different recombination algorithms, the resulting distributions might differ slightly. Testing CA and anti-k_T for the same pruning criteria leads to the results shown in figure 5.15. Both distributions, figure 5.15b and figure 5.15a, were obtained using the same approach as for figure 5.14 and show approximately the same results as the k_T algorithm.

5.2.4 Re-scaling Pruned Jet Masses

As mentioned earlier, pruning may change the jet properties such as the mass. In the case of small pruning, this will only slightly shift the jet masses. However, a stronger pruning, such as the one discussed above, may significantly shift the jet mass depending on the substructure of the jet. The impact of pruning using $D_{cut} = 0.05$ and $z_{cut} = 0.7$ is shown in figure 5.16. Here, the mass of the combined jet pairs are drawn. While the selected jet pairs are supposed to reconstruct the Higgs or Z boson, the masses are shifted to a much lower region due to pruning. Since information on the first daughter particles of each boson is accessible, this "truth" information may be used to select the jets corresponding to the quarks as well as for selecting the correct pairs. However, this information is not available for a measurement. To mimic the scenario of an experimental measurement, only a minimal amount of "truth"

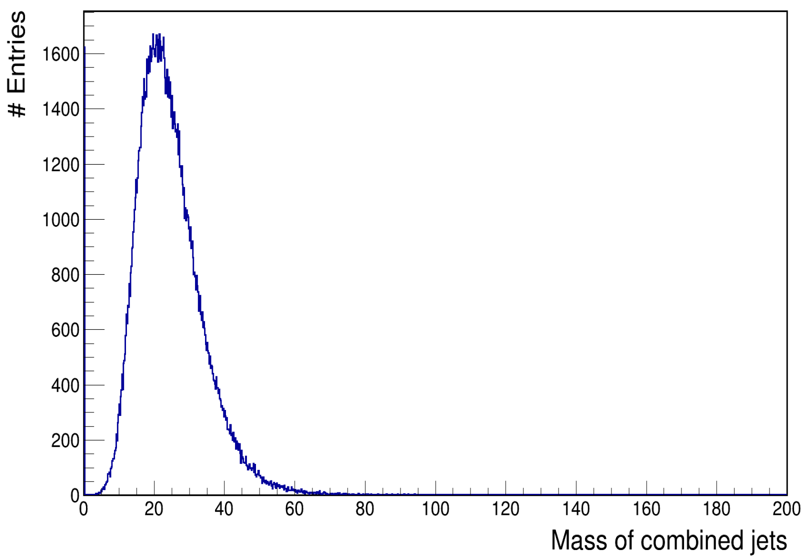


Figure 5.16: The distribution for the mass of the jet pairs reconstructing the Z boson. Here, pruning with $z_{cut} = 0.7$ and $D_{cut} = 0.05$ was applied to the jets. Additionally, $\Delta R \leq 0.5$ was used for the jet selection and "truth" information about the origin of the bottom quarks was used to select the correct jet pairs. While the mass of the reconstructed Z boson should be around 90 GeV, this distribution shows a significant shift to lower mass regions due to the applied pruning.

information should be used. While this work uses the spatial difference, which requires the four-vector of the bottom quarks, this is solely used to mimic b-tagging. Omitting any additional information on the origin of the jets would require a variable used to select the jet pairs. This work previously used the mass of the combined jets and compared them to the boson masses. Choosing the pair closest to the Higgs or Z boson in mass was then used to build the three-jet system. However, using the pruning specified above imposes a challenge on this selection method. To be able to further use the mass as a selection criterion for the jet pairs, the four-vector after pruning may be scaled in such a way that its mass is shifted to the correct mass window for Higgs and Z boson again. This scaled jet is then solely used for the selection of the jet pairs while the pruned jets are used in the computations.

The scaling of the jet mass after pruning may be applied to either the mass of the jet directly or to the components of the four-vector of the jet. Further, it is useful to investigate if an event-wise scaling is sufficient rather than applying a separate scaling for every jet in an event. Calculating a mass ratio for the corresponding pruned and non-pruned jets and multiplying the jet masses by this value is able to shift the masses to the correct region. However, scaling every jet mass by its respective mass ratio reproduces the mass distribution obtained by the non-pruned jets. This may be used to select the jet pairs and results in similar distributions compared to figure 5.17. This jet-wise scaling has the same effect as simply using the mass of the non-pruned jets for the selection. To investigate if it is possible to optimize scaling, the components of the four vectors of the jets may be scaled. For this purpose, ratios of the jet momenta and energies can be calculated for every pruned jet and its corresponding non-pruned version. Further, such a ratio may be calculated event-wise or jet-wise. For event-wise scaling, the mean of the momentum ratios and energy ratios needs to be computed and each component of the vector is multiplied by its respective scaling factor. While this may be able to shift the jet mass, there is no apparent improvement compared

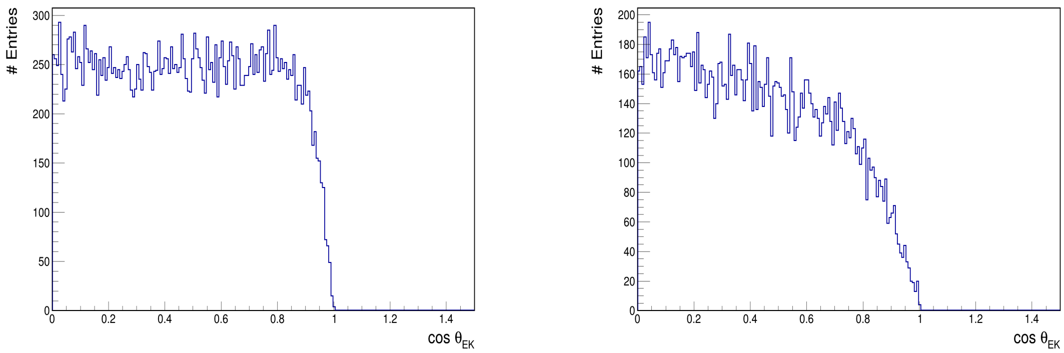
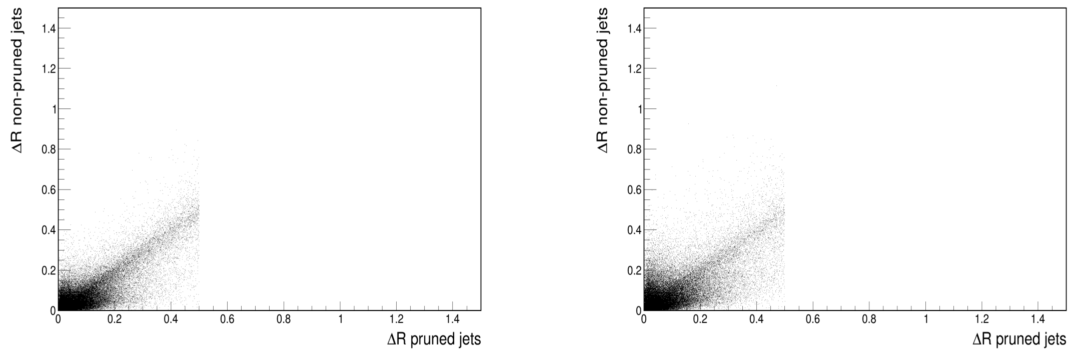


Figure 5.17: These distributions for $\cos\theta_{EK}$ were obtained by applying a pruning with $D_{cut} = 0.05$ and $z_{cut} = 0.7$ to the reconstructed jets. Additionally, the four final state jets were selected using $\Delta R \leq 0.5$. In order to select the jet pairs reconstructing the heavy bosons the jet mass was scaled to fit the mass window of the Higgs and Z boson. Figure 5.17a shows the distribution for the Higgs boson. Similar to previous results, it fits the expected "box" shape for a scalar particle. Figure 5.17b shows the distribution obtained for the Z boson, showing a significant improvement to previous results not applying pruning. Such results were previously obtained by using "truth" information for the jet selection or computing ROOTEK on particle level, showing that the scaled mass combined with pruning is able to significantly improve the jet selection.

to simply using the masses of the non-pruned jets. Jet-wise scaling is the most detailed scaling tested for this work. Here, every four-vector component is multiplied by scaling factors obtained by comparing the pruned jet to the non-pruned one. Here, the momenta of pruned and non-pruned jets give the resulting scaling factor for the momentum components. Additionally, the energy of the jets is compared, which provides a scaling factor for the energy component of the four-vector. This scaling produces a four-vector similar to the non-pruned jet and ultimately results in a mass distribution close to the non-pruned jet masses. This leads to results similar to using the mass of the non-pruned jets to select the correct pairs reconstructing the Higgs and Z bosons. Figure 5.17 shows the distribution resulting from applying the strong pruning specified earlier as well as $\Delta R \leq 0.5$ and the scaled mass. Here, the pruned version of the jets is used to test the ΔR threshold and select four jets which correspond to the bottom quarks. These jets are afterwards combined in every possible way and their scaled mass is computed, which is then used to select the combination that fits the hypothesis best. Afterwards, the pruned jets corresponding to the selected jet pair are used for the computation of ROOTEK. Figure 5.17 shows that this method is able to improve the distribution for the Z boson significantly, producing results similar to figure 5.14, which uses the available information on the origin of the quarks.

5.2.5 Study of Larger R Jets

While pruning the existing jets and scaling the jet masses provides significant improvements to the obtained results for the Z boson, there is still room for further improvement of the distribution of $\cos\theta_{EK}$ obtained by ROOTEK. This work so far applied a jet reconstruction algorithm with an R parameter of $R = 0.4$. However, there are analyses using large-R jets to identify the jets reconstructing the Higgs boson.(40) Using a larger R parameter for the initial jet reconstruction may open the opportunity for even stronger pruning without risking



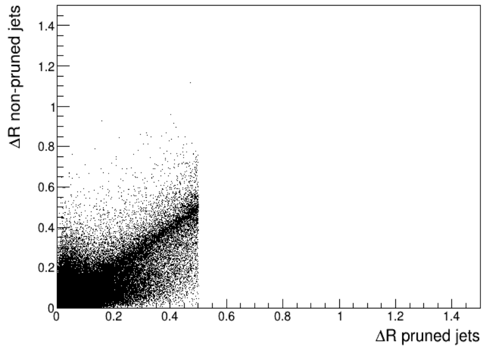
(a) (b)

Figure 5.18: The distributions shown draw ΔR for the pruned jets on the x-axis and ΔR for the corresponding non-pruned jets on the y-axis. To obtain figure 5.18a, pruning with $z_{cut} = 0.7$ and $D_{cut} = 0.00005$ is applied showing clustering at small values of ΔR. Figure 5.18b results from using pruning with $z_{cut} = 0.9$ and $D_{cut} = 0.00005$ leading to an even stronger clustering at small values of ΔR.

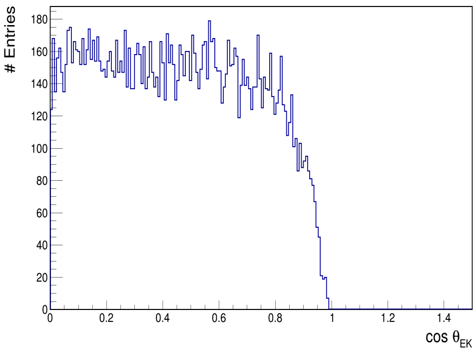
lower statistics. To investigate this effect, $R = 1$ defining a larger angular reach is tested. Comparing the same pruning applied to $R = 0.4$ and $R = 1$ jets shows an increase in the statistics found for $\Delta R \leq 0.5$. Figure 5.18 presents two prunings of different strengths applied on large-R jets. Here, pruning with lower D_{cut} was used. Additionally, figure 5.18a applied $z_{cut} = 0.7$ and figure 5.18b $z_{cut} = 0.9$. Both show clustering at low ΔR values, indicating a better jet selection achieved by strong pruning. Further, comparing both distributions visualizes higher statistics in figure 5.18b, which used stronger pruning. This shows that using large-R jets may improve the jet selection. However, this needs to be investigated further.

As an indication of the possible impact of large-R jets, the pruning used for the previous results is applied in addition to "truth" information on the origin of the quarks. This results in the distribution shown in figure 5.19a for ΔR of pruned and non-pruned jets. This figure shows a much larger number of entries in the region $\Delta R \leq 0.5$ than previously, resulting from the larger angular reach of the initial jet reconstruction. Additionally, the majority of the entries are clustered around small values of ΔR. This will ultimately improve the selection of the four jets corresponding to the quarks. Afterwards, the "truth" information is applied instead of using the scaled mass to select the jet pairs reconstructing the heavy bosons. Ultimately, the obtained distribution of $\cos\theta_{EK}$ in figure 5.19c for the Z boson is improved compared to the initial results.

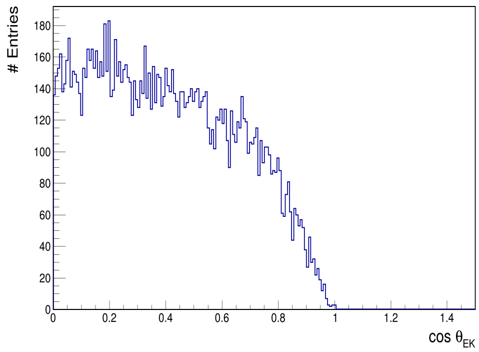
Even though the distributions of $\cos\theta_{EK}$ can be improved significantly by using a better jet selection, further improvements are required. Another method for the selection of the four final state jets would make use of the jet constituent with the highest transverse momentum. For this purpose, the constituents of the jet are iterated to find the contribution with the highest p_T. The four-vector of this is then compared to the quarks in the same way as before. Then, either "truth" information or the scaled mass may be used to select the jet pairs. However, further investigations on how to correctly combine pruning with the highest p_T constituent of the jets need to be done to achieve an improved clustering towards low values of ΔR.



(a)



(b)



(c)

Figure 5.19: This distribution in figure 5.19a shows the spatial difference of jets and quarks using large-R jets. On the x-axis, ΔR of the pruned jets is drawn and on the y-axis, ΔR of the corresponding non-pruned jets. Here, the R parameter for the initial jet reconstruction is set to $R = 1$. Additionally, pruning is applied with $D_{cut} = 0.05$ and $z_{cut} = 0.7$. Further, "truth" information on the origin of the quarks is used to select the jet pairs. This distribution shows large clustering towards lower values of ΔR, which may ultimately improve the jet selection. Figure 5.19b and figure 5.19c show the results achieved for $\cos\theta_{EK}$ using large-R jets in the initial jet reconstruction. Further, the same selection criteria as for figure 5.19a are applied. Figure 5.19b shows the results for the Higgs boson and figure 5.19c for the Z boson, which shows an improvement compared to the initial results.

5.2.6 Conclusion

While even the initial distribution for the Higgs boson showed the "box"-shaped distribution for $\cos\theta_{EK}$ as expected for a scalar particle, the initial results for the Z boson did not fit the decreasing distribution expected for a spin-1 particle. Due to this, the jets used to reconstruct the boson were investigated using different approaches. These proposed approaches led to an improvement of the jet selection, which then ultimately improved the resulting distribution of $\cos\theta_{EK}$. The work in this section shows that the four final state jets identification needs to be precise to be able to distinguish between the Higgs and Z bosons in the HZ processes. If this selection is sufficiently good, the jet pairs can then be selected by using mass intervals for the Higgs and Z bosons.

5.3 Separation of HH and HZ Processes Using a Neural Network

While the previous analyses of the HH and HZ processes provided methods to improve ROOTEK to enable distinguishing between Higgs and Z bosons using this observable, other kinematic quantities remain very similar for both massive bosons. Due to this, the separation of these processes requires an algorithm able to capture details of various distributions obtained for the Higgs and Z bosons. Such an information-processing algorithm is provided by Neural Networks introduced in chapter 3 which is implemented in Python 3.7 using a plugin called PyTorch (41). Before the actual classification of final states using a NN, variables able to distinguish HH and HZ final states need to be found. This chapter presents some of the variables hinting at useful characteristics to become included in a network data set. Additionally, the achieved accuracy for the classification of HH and HZ final states is discussed and the impact of the investigated spin correlation is evaluated.

The variables used in a network data set can be grouped into properties of the reconstructed bosons and properties of the final state jets. As introduced in the previous chapter 5.1, the jet masses are used to identify the possible reconstructed Higgs and Z bosons. For this purpose, all possible combinations are built and compared to the boson masses. Further, in the analysis of HZ processes, an additional selection criterion ΔR is included in the jet selection. This is used to improve the jet selection. Further, it ensures selecting jets that most probably originated from bottom quarks used to reconstruct the Higgs and Z bosons instead of randomly combined jet pairs. While the mass distribution of the Higgs and Z bosons overlap slightly, the analysis presented in this work uses separate mass intervals. Due to this, the masses of the selected jet pairs may be a useful quantity to distinguish between Higgs and Z bosons. Additionally, the masses of the corresponding pruned jet pairs may be used. Even though there is a significant overlap between the mass distributions of the pruned jet pairs for Higgs and Z bosons, there are some cases for the reconstructed Higgs boson, where the pruned jet pairs have a mass around 125 GeV. Additionally, pruned jets were used for the jet selection, meaning that their properties may provide further useful characteristics to distinguish between the final states. Further, slight differences in the transverse momentum, which is defined in the x-y plane, as well as the energy of Higgs and Z bosons may provide additional details for the classification and provide a measure to distinguish between HH and HZ final states. The angular coordinates Φ and η of the reconstructed bosons may be used. However, these variables prove to be more useful when investigated between the jet pairs. While $\Delta\eta$ between the combined jets shows an immense overlap, the difference ΔR, as well as $\Delta\Phi$, show significant differences for the jet pairs of the Higgs and Z boson.

Figure 5.20 presents the distributions of $\Delta\Phi$ for the Higgs and Z boson in the HZ processes. To obtain these distributions the bosons are reconstructed using "truth" information on the origin of the quarks. While both show peaks at about $\Delta\Phi = \pm 4$, the results for the Z boson in figure 5.20b show more narrow peaks and additional side maxima close to 0. These significant differences result in ΔR differing for the Higgs and Z bosons. This makes both angular differences a useful property to become included in a network data set. However, using ΔR does not lead to improved results. While most variables will not show a significant difference between the Higgs and Z bosons, the obtained distributions for inaccurate jet pairs randomly having the correct mass may differ largely, making these variables good indicators of a correct jet pair selection.

All previously mentioned variables may be written into the network data sets for HH and HZ

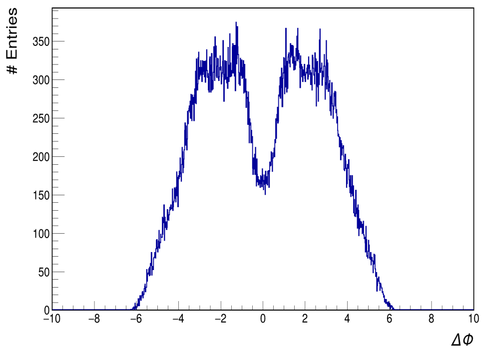


Figure 5.20: The distributions for $\Delta\Phi$, measured between the jet pairs reconstructing the Higgs and Z bosons. Figure 5.20a shows the distribution resulting for the Higgs boson and figure 5.20b for the Z boson. While both show peaks in the same area, the peaks for the jet pairs of the Z boson are more narrow.

processes in addition to $\cos\theta_{EK}$. These are then fed to a NN with one hidden layer. The error function used is the cross-entropy loss briefly introduced in chapter 3 with a ReLu activation in the hidden layer and softmax activation in the output layer. Additionally, the input data is normalized using the mean and the standard deviation. As mentioned in chapter 3, the input data set contains a column specifying the target values and is split into a training and validation subset. Lastly, a batch size of 600 and 150 epochs are used.

The data set which is given to the NN was generated for the strongly pruned jets. Additionally, a threshold of $\Delta R \leq 0.5$ and mass scaling are applied. As shown in figure 5.21, the network was able to correctly classify the HH and HZ final states for about 76% of the data points, which indicates a good performance of the network. However, the data points were selected by applying several restrictions. Firstly, the Monte Carlo data with initially 1.2 million events is limited to events consisting of at least four jets. Afterwards, these jets are pruned and a ΔR criterion is applied. If for all four bottom quarks, there is a jet that fits the threshold, the event is used, which further limits the data set. Lastly, the mass of the jet pairs is tested. Since the jet masses lie on a distribution with a certain width, there may be events which do not contain jet pairs with a mass fitting in the selected mass intervals. However, $\cos\theta_{EK}$ is obtained by ROOTEK, which does not need phase-space restrictions. These applied criteria result in a selection of events which might show an idealized version of the obtainable distributions. This may impact the performance accuracy of the NN, which means that a value of 76% might not be achievable in a less restricted data set. Additionally, all of the previously mentioned selection criteria reduce the amount of data points. Chapter 3 describes that NNs need to process as much unfamiliar information as possible to ensure a good performance. While the number of batches iterated over in an epoch or the number of epochs can be changed, the information obtainable from the network data set is limited by its number of data points. The evaluation of the NN's results for different data sets shows that the performance of the network is sensitive to the size of the input data set. While the input data set for the results of figure 5.17 achieves a classification accuracy of 76% a case with fewer data points due to a different pruning achieves a far lower accuracy. Currently, these results hint at a sensitivity to the size of the data set but this needs to be investigated further. Further, the NN can be tested with data sets not containing $\cos\theta_{EK}$ which may indicate the impact of this variable on the classification accuracy. However, current results do not show a difference in the resulting accuracy. Since there are no immense differences

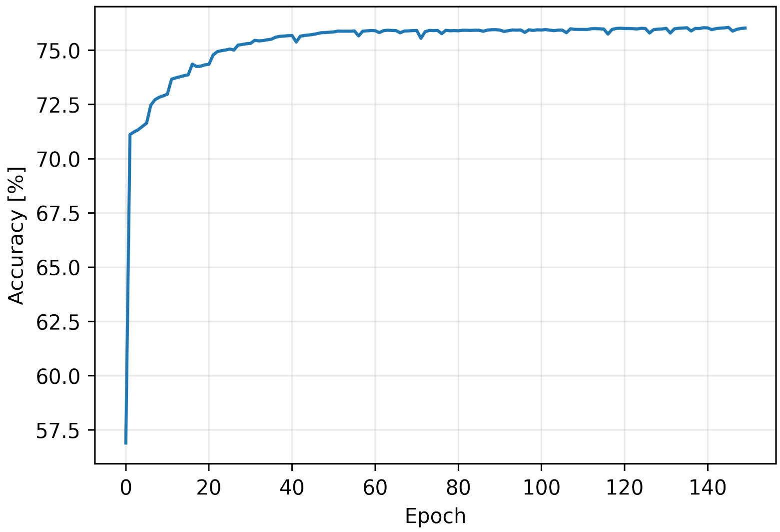


Figure 5.21: The classification accuracy is drawn as a function of the number of epochs. The model used is able to correctly classify HH and HZ final states for 76% of the data points of the selected events.

in the distributions of Higgs and Z bosons for this variable, only minor differences would be expected. The impact of $\cos\theta_{EK}$ on the learning of the network needs to be studied further. For instance, distributions obtained using "truth" information might hint at the best achievable separation of HH and HZ processes using $\cos\theta_{EK}$. However, first tests showed the NN not learning properly using the "truth" information for the jet pair selection. Further tests to evaluate the potential of $\cos\theta_{EK}$ might use the scaled mass for the jet pair selection and use "truth" information only for $\cos\theta_{EK}$. Here, challenges may arise from inaccurate correlations between the input variables which arise from the different methods to obtain the distributions. If it is possible to find an approach which can mimic the best achievable separation in $\cos\theta_{EK}$ for the Higgs and Z bosons which can be tested in a NN, the maximum allowed error in the jet pair selection can be tested by gradually increasing the error of this selection and monitoring the performance of the NN. Additionally, the performance of the network may be improved by increasing the statistics of the input data and investigating further possible variables which can be used to distinguish the Higgs and Z bosons.

Overall a good separation between HH and HZ final states can be achieved for the selected events including $\cos\theta_{EK}$. However, further testing and improvements are necessary to be able to evaluate the usability for analyses in the field of HH production.

Chapter 6

Summary

This work investigated the impact of the spin on the direction of the final state jets for HH and HZ processes. For this purpose, two approaches were introduced: MODEK and ROOTEK. While both should in general lead to the same results, results obtained by MODEK contained fewer entries in every distribution. Investigating subsets of the phase-space showed a better agreement between MODEK and ROOTEK for large pseudorapidity regions, $\eta \geq 2.5$. However, there were remaining disagreements between MODEK and ROOTEK. Further investigations presented perfect agreements for variables such as the boost velocity β used in both approaches which hinted at issues within the derived expression for MODEK. For certain cases, non-defined values for $\cos\theta_{EK}$ arise from the square-root term in MODEK. Since ROOTEK produced reliable results for the HH processes, the investigation of the HZ processes focused on this approach. Initial results did not present the expected distribution for the Z boson. However, calculating $\cos\theta_{EK}$ on the particle level showed that the necessary information about the particle spin is available in the generated data. In order to improve the distribution for the spin-1 particle, various methods to improve the jet selection are applied. Ultimately, a combination of a ΔR threshold between jets and quarks, strong jet pruning and mass scaling provided the best results for the Z boson. This shows that a precise jet selection is necessary to be able to distinguish between Higgs and Z bosons using $\cos\theta_{EK}$. Additional tests to improve the jet selection were briefly introduced which need to be investigated further. Lastly, the achievable classification accuracy was obtained using a NN. This shows that a good classification of HH and HZ final states is achievable for the selected events.

Future work could focus on the investigation of the jet selection to evaluate the precision which is needed to use $\cos\theta_{EK}$ to distinguish between HH and HZ final states. Additionally, further tests on the impact of $\cos\theta_{EK}$ on the performance of the NN can be performed. Additionally, the expression obtained for MODEK can be investigated further in order to find an equation which may be usable in likelihood fits. Further, the tests so far were conducted on purely simulated data without any noise or background processes, which means, that additional work to test the presented approach for these circumstances is required.

In general, it was presented that the used approach provides a variable which can be used to distinguish between HH and HZ final states with a sufficient jet selection. However, both approaches tested in this work require further improvement to make the usage for an analysis possible.

Bibliography

[1] N. A. Bahcall, *Dark matter universe*, Proceedings of the National Academy of Sciences **112** no. 40, (2015) 12243–12245, https://www.pnas.org/doi/pdf/10.1073/pnas.1516944112. https://www.pnas.org/doi/abs/10.1073/pnas.1516944112.

[2] D. Farrah, K. S. Croker, M. Zevin, G. Tarlé , V. Faraoni, S. Petty, J. Afonso, N. Fernandez, K. A. Nishimura, C. Pearson, L. Wang, D. L. Clements, A. Efstathiou, E. Hatziminaoglou, M. Lacy, C. McPartland, L. K. Pitchford, N. Sakai, and J. Weiner, *Observational Evidence for Cosmological Coupling of Black Holes and its Implications for an Astrophysical Source of Dark Energy*, The Astrophysical Journal Letters **944** no. 2, (2023) L31. https://doi.org/10.3847

[3] M. Císcar-Monsalvatje, A. Ibarra, and J. Vandecasteele, *Matter-antimatter asymmetry and dark matter stability from baryon number conservation*, 2023. https://doi.org/10.48550/arXiv.2307.02592.

[4] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** no. 1, (2012) 1–29. https://www.sciencedirect.com/science/article/pii/S037026931200857X.

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** no. 1, (2012) 30–61. https://www.sciencedirect.com/science/article/pii/S0370269312008581.

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509. https://doi.org/10.1103/PhysRevLett.13.508.

[7] J. Ellis and I. Karliner, *Measuring the spin of the gluon in* e^+e^- *-annihilation*, Nuclear Physics B **148** no. 1, (1979) 141–147. https://www.sciencedirect.com/science/article/pii/0550321379900191.

[8] C. B. et al., *First observation of hadron production in* e^+e^- *collisions at 13 and 17 GeV CMS energy with the pluto detector at petra: Pluto Collaboration*, Physics Letters B **81** no. 3, (1979) 410–415. https://www.sciencedirect.com/science/article/pii/0370269379903654.

[9] V. Gribov, *The theory of quark confinement*, The European Physical Journal C **10** no. 1, (1999) 91–105. https://doi.org/10.1007

[10] Particle Data Group Collaboration, R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01. https://doi.org/10.1093/ptep/ptac097.

[11] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008. https://cds.cern.ch/record/111880.

[12] M. A. et al., *Measurement of the W-boson mass in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, The European Physical Journal C **78** no. 2, (2018). https://doi.org/10.1140

[13] L. Di Lella and C. Rubbia, *The Discovery of the W and Z Particles*, Adv. Ser. Direct. High Energy Phys. **23** (2015) 137–163. https://doi.org/10.1142/9789814644150$_0$006.

[14] W. N. Cottingham and D. A. Greenwood, *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 2 ed., 2007. https://doi.org/10.1017/CBO9780511791406.

[15] A. Bettini, *Introduction to Elementary Particle Physics*. Cambridge University Press, 2008. https://doi.org/10.1017/CBO9780511809019.

[16] E. Bagnaschi, H. Bahl, E. Fuchs, T. Hahn, S. Heinemeyer, S. Liebler, S. Patel, P. Slavich, T. Stefaniak, C. E. M. Wagner, and G. Weiglein, *MSSM Higgs boson searches at the LHC: benchmark scenarios for Run 2 and beyond, journal = The European Physical Journal C,*. https://doi.org/10.1140

[17] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*. Cambridge University Press, 3 ed., 2018. https://doi.org/10.1017/9781316995433.

[18] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3*, 2023. https://doi.org/10.48550/arXiv.2305.16623.

[19] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** no. 08, (2008) S08001. https://dx.doi.org/10.1088/1748-0221/3/08/S08001.

[20] B. Schmidt, *The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments*, Journal of Physics: Conference Series **706** no. 2, (2016) 022002. https://dx.doi.org/10.1088/1742-6596/706/2/022002.

[21] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** no. 08, (2008) S08003. https://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[22] B. Mehlig, *Machine Learning with Neural Networks: An Introduction for Scientists and Engineers*. Cambridge University Press, 2021. https://doi.org/10.1017/9781108860604.

[23] I. Sarker, *Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions*, SN Computer Science **2** (2021). https://doi.org/10.1007/s42979-021-00815-1.

[24] P. Cunningham and S. J. Delany, *Underestimation Bias and Underfitting in Machine Learning*, pp. , 20–31. Springer International Publishing, Cham, 2021.

[25] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, Journal of High Energy Physics **2010** no. 6, (2010). https://doi.org/10.1007

[26] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, T. Sjöstrand, P. Skands, M. Utheim, and R. Verheyen, *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2022. https://doi.org/10.48550/arXiv.2203.11601.

[27] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, The European Physical Journal C **72** no. 3, (2012). https://doi.org/10.1140

[28] A. Buckley and D. B. Gupta, *Powheg-Pythia matching scheme effects in NLO simulation of dijet events*, 2017. https://doi.org/10.48550/arXiv.2203.11601.

[29] R. Brandelik et al., *Evidence for a spin-1 gluon in three-jet events*, Physics Letters B **97** no. 3, (1980) 453–458. https://www.sciencedirect.com/science/article/pii/0370269380906395.

[30] SLD Collaboration, C.-G. Fan, *Measurements of gluon spin sensitive quantities at the Z^0 resonance*, pp. , 898–900. 1992.

[31] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** no. 1, (1997) 81–86. https://www.sciencedirect.com/science/article/pii/S016890029700048X. New Computing Techniques in Physics Research V.

[32] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, 2009. https://doi.org/10.48550/arXiv.0901.0512.

[33] M. Casado, *A High-Granularity Timing Detector for the ATLAS Phase-II upgrade*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1032** (2022) 166628. https://www.sciencedirect.com/science/article/pii/S0168900222002005.

[34] ATLAS Collaboration, G. Piacquadio, *B-tagging at LHC: Expected performance and its calibration in data*, PoS **CHARGED2008** (2008) 033. https://doi.org/10.22323/1.073.0033.

[35] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, *Towards an understanding of jet substructure*, Journal of High Energy Physics no. 9, (2013). https://doi.org/10.1007

[36] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Physical Review D **80** no. 5, (2009). https://doi.org/10.1103

[37] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches*, Physical Review D **81** no. 9, (2010). https://doi.org/10.1103

[38] CMS Collaboration, *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging*,. https://cds.cern.ch/record/1194489/files/JME-09-001-pas.pdf.

[39] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, Journal of High Energy Physics no. 04, (2008) 063–063. http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[40] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider*, Physical Review Letters **100** no. 24, (2008). https://doi.org/10.1103

[41] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, 2019. https://doi.org/10.48550/arXiv.1912.01703.

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Arbeit mit dem Titel

Separation of HH and HZ Final States Using Spin Correlations

selbständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Celine Stauch

München, den 27. November 2023